University of Alberta

Molecular Investigation of a Retinal Disease Hotspot on Human Chromosome 6

by

Pamela Sarita Lagali ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Molecular Biology and Genetics

Department of Biological Sciences

Edmonton, Alberta

Spring 2004

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-96290-3
Our file Notre référence
ISBN: 0-612-96290-3

The author has granted a non-exclusive license allowing the Library and Archives Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

Retinal degeneration characterizes a broad spectrum of ocular disorders and is the leading cause of blindness in the developed world. Deterioration of the central or macular region of the retina is responsible for the majority of cases of severe vision loss. A form of autosomal dominant macular dystrophy segregating in a large Canadian pedigree has been identified. The clinical features of the disease resemble other documented forms of macular degeneration. Genetic analysis of the family revealed that the disease locus lies on the long arm of human chromosome 6, a genomic region to which several other retinal disease loci have been mapped. As a resource to define candidate disease genes, a retinal and retinal pigment epithelial transcript map was developed for the critical chromosome 6q interval harbouring the genetic defect. The *C6orf37* gene was identified as a retina-expressed, evolutionarily conserved gene that resides within the critical interval, however mutational analysis excluded it as the disease-causing gene within the Canadian pedigree. Isolation of the *ELOVL4* gene, also found within the 6q interval, and subsequent mutation screening revealed it to be the disease gene for a number of ancestrally related kindreds including the Canadian family. This novel gene was further characterized as a member of a highly conserved gene family that appears to be involved in fatty acid biosynthesis. Mutagenesis of the encoded protein and subcellular localization studies demonstrate that macular disease-associated mutations lead to disrupted intracellular trafficking of the ELOVL4 protein. Additional mutant forms of ELOVL4 have been generated and are currently being analyzed towards functional dissection of the protein. These studies provide molecular insights into pathogenic mechanisms and potential therapeutic strategies for retinal degenerative conditions.

"Knowledge is the true organ of sight, not the eyes."

Panchatantra (c. 5th century A.D.)
Translation from Sanskrit by Franklin Edgerton

This thesis is dedicated to
Paul Wong
who took a leap of faith
and made all the difference.

ACKNOWLEDGEMENTS

There are many people to whom I owe a debt of gratitude for their role in helping this thesis and doctoral degree come to fruition.

I don't know how to even begin to thank my supervisor, Paul Wong, for what he has done to so profoundly change my life. From taking a chance when no one else would, to being a true mentor in every sense of the word, I have never seen such dedication, commitment, and self-sacrifice towards the training, development, and success of students. Paul, you have taught me so much, opened so many doors for me, and created a path for me to realize my academic goals. No amount of thanks could ever be enough for all that I have learned and gained from you during the past five and a half years of the fullest graduate experience I could have hoped for. It has been such a privilege to work with you, and I look forward to our continued collaboration and friendship for many years to come.

Since virtually the start of my doctoral program, I have worked very closely with and attribute much of my progress to the efforts of Dr. Radha Ayyagari and her laboratory associates, in particular Dr. Paul Sieving, Dr. Irina Griesinger, Laura Kakuk, Dr. Jiafan Liu, and Dr. Rajesh Ambasudhan, at the W.K. Kellogg Eye Center at the University of Michigan. Radha, I am sincerely grateful to have had the opportunity to come to your lab to learn the techniques of human genetic analysis, and to participate in our productive collaboration since that time. I have also very much enjoyed both our scientific and personal discussions over the years. You are a true role model for me as a successful woman in science and particularly in the field of vision research, and I hope to one day be able to juggle it all even half as well as you do.

I greatly appreciate the contributions of Dr. Ian MacDonald to the initiation and progression of our study of the Canadian STGD3 pedigree. Thanks also to both Ian and Dr. Heather McDermid for the guidance and support they have provided as members of my graduate supervisory committee throughout my doctoral program. I thank the additional members of my candidacy and thesis defense examining committees for their time and efforts in enabling me to reach this stage. A special note of gratitude is extended to Dr. William Samuel, Dr. Carol Cass, and Dr. Roseline Godbout for their supportive role in initiating the whole process.

My time in the Wong Lab has been enriched by the intellectual stimulation and camaraderie provided by the numerous students that have passed through its doors since my arrival. The graduate students - Michelle Patterson, Iga Stepczynski, Ben McDonald, Ruby Grewal, Rhonda Kelln, Micah Chrenek, Brian Maciejko, Dan Sprague, and Alison Ziesel - and the many undergraduate students, notably Disa Brownfield, Chris Carlson, Denise Patterson, Greg Lutzak, and Tim Erickson, have contributed to the projects I have worked on, have made the lab an always fun place to work, and have created many memorable moments for me both in and out of the lab (especially during our many ARVO adventures!). Thank you all for your valuable help and friendship.

I would like to acknowledge the assistance provided by Dr. Steven Bernstein, Dr. T. Michael Redmond, and Dr. Diane Borst in various projects with which I have been involved, and also for our interactions at numerous conferences, making them all the more enjoyable experiences.

I am very fortunate to have the constant support and encouragement of my parents, Kokila and Sudhakar Lagali, and my brother and sister-in-law, Neil and Theresa Lagali. Their help, words of wisdom, and love have carried me through the challenging times and have made the successes more special, and I never would have made it to this point without them. I am also thankful to the many friends I have made during my stay in Edmonton who have made it a real home and have become an extended family to me.

Finally, I would like to express heartfelt gratitude to my husband, Miguel Cabrita, for his assistance at every stage of this process. The continual sharing of intellectual and technical expertise, contribution of various reagents and protocols, and reading and editing of manuscripts, including this thesis, have been instrumental in my achievements throughout my graduate career. In addition, his unwavering belief in me, unselfish love and patience, and enthusiastic support of my chosen path at every turn have been an ongoing source of inspiration that I truly cherish.

The entirety of the work presented in this thesis has been generously funded by a number of organizations including the Alberta Heritage Foundation for Medical Research, the E.A. Baker Foundation of the Canadian National Institute for the Blind, The Canadian Institutes of Health Research, the Foundation Fighting Blindness-Canada, and the University of Alberta, to all of which I am extremely grateful.

TABLE OF CONTENTS

Chapter **Page**

1. INTRODUCTION 1

1.1 The Human Eye 2
1.2 The Retina 3
1.2.1 Retinal Cell Types and Spatial Organization 3
1.2.1.1 Photoreceptors 4
1.2.1.2 Retinal Pigment Epithelium 5
1.2.2 Phototransduction 5
1.3 Inherited Diseases of the Retina 6
1.3.1 Diagnosis of Retinal Degenerative Phenotypes 7
1.3.1.1 Psychophysical Parameters 7
1.3.1.2 Fundus Imaging 8
1.3.1.3 Fluorescein Angiography 8
1.3.1.4 Electroretinography 9
1.3.2 Regional Variation of Retinal Degeneration 10
1.3.3 Genetic Diversity of Retinal Degeneration 10
1.3.4 Age-Related Macular Degeneration 11
1.4 Focus of the Thesis 12
1.5 References 13

2. AUTOSOMAL DOMINANT STARGARDT-LIKE MACULAR DYSTROPHY SEGREGATING IN A LARGE CANADIAN FAMILY 30

2.1 Introduction 31
2.2 Materials and Methods 32
2.2.1 Subjects 32
2.2.2 Genomic DNA Isolation 33
2.2.3 Genotyping 33
2.2.4 Linkage Analysis 34
2.2.5 Haplotype Analysis 34
2.3 Results 34
2.3.1 Clinical Examination 34
2.3.2 Genetic Analysis 35
2.4 Discussion 36
2.5 Acknowledgements 39
2.6 References 39

3. A TISSUE-SELECTIVE TRANSCRIPT MAP FOR A RETINAL DISEASE LOCUS-RICH REGION OF HUMAN CHROMOSOME 6q 56

3.1 Introduction 57
3.2 Materials and Methods 60
3.2.1 Defining the Chromosome 6q12-q16 Genomic Interval of Interest 60

3.2.2 Identification and Expression Analysis of ESTs 60
3.2.3 Construction of the Retinal and RPE Transcript Map 61
3.2.4 Expression Profiling and Evolutionary Conservation of Putative Retinal/RPE Genes 61
3.3 Results 62
3.3.1 Identification of Retina- or RPE-Expressed 6q12-q16 ESTs 62
3.3.2 Categorization of Ocular Genes Represented by 6q12-q16 ESTs 63
3.3.3 A Comprehensive Retinal and RPE Transcript Map on 6q12-q16 64
3.3.4 Further Characterization of Retina/RPE-Expressed 6q12-q16 Genes 65

3.4 Discussion 66
3.5 Acknowledgements 69
3.6 References 69

4. IDENTIFICATION AND CHARACTERIZATION OF *C6orf37*, A NOVEL CANDIDATE HUMAN RETINAL DISEASE GENE ON CHROMOSOME 6q14 94

4.1 Introduction 95
4.2 Materials and Methods 95
4.2.1 Identification and Analysis of ESTs Localized to 6q14-q16 95
4.2.2 Isolation of C6orf37 cDNA 96
4.2.3 C6orf37 DNA and Protein Sequence Analysis 97
4.2.4 Zooblot Analysis 97
4.2.5 Expression Analysis by Northern Blotting 98
4.2.6 *In vitro* Translation 98
4.2.7 Mutational Analysis of *C6orf37* in STGD3/adMD Pedigrees 98
4.3 Results 99
4.3.1 Isolation of Human C6orf37 cDNA 99
4.3.2 Determination of *C6orf37* Genomic Structure 100
4.3.3 Sequence Conservation of C6orf37 101
4.3.4 C6orf37 Protein and mRNA Sequence Analysis 101
4.3.5 Tissue Distribution of C6orf37 mRNA 102
4.3.6 Exclusion of *C6orf37* as the STGD3/adMD Causative Gene 102
4.4 Discussion 103
4.5 Acknowledgements 105
4.6 References 105

5. A 5-bp DELETION IN THE NOVEL RETINAL GENE *ELOVL4* IS ASSOCIATED WITH RELATED FORMS OF AUTOSOMAL DOMINANT MACULAR DYSTROPHY 121

5.1 Introduction 122
5.2 Materials and Methods 123
5.2.1 Family Material and Linkage Analysis 123
5.2.2 ELOVL4 cDNA Cloning 124
5.2.3 Organization of the *ELOVL4* Gene 125
5.2.4 Bioinformatic Protein Sequence Analysis 126
5.2.5 RNA Isolation and Northern Blot Analysis 126

    5.2.6 *ELOVL4* Mutation Analysis 127
5.3 Results 128
    5.3.1 Defining a 0.6-cM Minimal Critical Interval for STGD3/adMD 128
    5.3.2 Identification of the *ELOVL4* Gene 128
    5.3.3 Characterization of *ELOVL4* Gene Products 129
    5.3.4 Identification of an *ELOVL4* Mutation in STGD3/adMD Patients 130
5.4 Discussion 131
5.5 Acknowledgements 136
5.6 References 137

6. EVOLUTIONARILY CONSERVED *ELOVL4* GENE EXPRESSION IN THE VERTEBRATE RETINA 155

6.1 Introduction 156
6.2 Materials and Methods 158
    6.2.1 Animal Tissues and Donor Eyes 158
    6.2.2 Southern Blot Analysis 158
    6.2.3 RNA Isolation and Northern Blot Analysis 159
    6.2.4 Production of ELOVL4 Antibodies 160
    6.2.5 Preparation of Retinal Protein Extracts 160
    6.2.6 SDS-PAGE and Immunoblot Analysis 160
    6.2.7 Bioinformatic Analysis 161
    6.2.8 Immunohistochemistry 161
        6.2.8.1 Human Retina Sections 161
        6.2.8.2 Animal Retina Sections 162
6.3 Results 162
    6.3.1 Multi-species Conservation of *ELOVL4* DNA Sequences and mRNA Expression 162
    6.3.2 ELOVL4 Protein Orthologues and Homologues 163
    6.3.3 Immunodetection of ELOVL4 Protein in Mammalian Retinas 165
    6.3.4 Distribution of ELOVL4 Protein in Mammalian Retinas 165
6.4 Discussion 166
6.5 Acknowledgements 169
6.6 References 170

7. FUNCTIONAL ANALYSIS OF THE ELOVL4 PROTEIN BY MUTAGENESIS AND SUBCELLULAR LOCALIZATION STUDIES 189

7.1 Introduction 190
7.2 Materials and Methods 192
    7.2.1 ELOVL4 Protein Sequence Analysis 192
    7.2.2 ELOVL4 cDNA Constructs and Site-Directed Mutagenesis 192
    7.2.3 Mammalian Cell Culture and Transfection 194
    7.2.4 SDS-PAGE and Immunoblotting 194
    7.2.5 Immunofluorescence Labeling and Microscopy 195
7.3 Results 195
    7.3.1 Mutagenesis of Putative Functional Domains in ELOVL4 195
    7.3.2 Expression of GFP-ELOVL4 Fusion Proteins in COS-7 Cells 196
    7.3.3 Intracellular Distribution of Wild-type ELOVL4 in COS-7 Cells 197

7.3.4 Subcellular Localization of Mutant ELOVL4 Proteins Carrying Macular Disease-Causing Deletions 197
7.3.5 ER Localization of ELOVL4 Depends on the Carboxy-Terminal Dilysine Motif 198
7.4 Discussion 198
7.5 Acknowledgements 202
7.6 References 203

8. SUMMARY AND CONCLUSION 221

8.1 Summary and Conclusion 222
8.2 References 226

LIST OF TABLES

Table **Page**

Table 2.1 Two-point LOD scores for linkage between the autosomal dominant Stargardt-like macular dystrophy locus within the Canadian pedigree and microsatellite markers on chromosome 6q 53

Table 3.1 A retinal and RPE transcript map for a segment of human chromosome 6q 78-87

Table 5.1 Primer pairs used for PCR amplification of *ELOVL4* exons for mutation analysis in STGD3/adMD families 152

LIST OF FIGURES

Figure		Page
Figure 1.1	Schematic diagram of the human eye	19
Figure 1.2	Structure of the retina	21
Figure 1.3	Geography of the macular and foveal regions of the human retina	23
Figure 1.4	Schematic diagram of the retinal pigment epithelium	25
Figure 1.5	The vertebrate visual phototransduction cascade	27
Figure 1.6	Funduscopy of the retina	29
Figure 2.1	Pedigree for the Canadian family showing autosomal dominant inheritance of a Stargardt-like macular dystrophy	45
Figure 2.2	Funduscopy and fluorescein angiography for two affected individuals within the Canadian autosomal dominant Stargardt-like macular dystrophy pedigree	47
Figure 2.3	Summary of visual acuity and colour vision tests for different individuals within the Canadian autosomal dominant Stargardt-like macular dystrophy family	49
Figure 2.4	Chromosome 6q microsatellite marker haplotypes for individuals within the Canadian autosomal dominant Stargardt-like macular dystrophy pedigree	51
Figure 2.5	Chromosome 6 ideogram showing the relative positions of the retinal disease loci that have been mapped to 6q	55
Figure 3.1	Dot blot analysis of chromosome 6q12-q16 ESTs for determining retinal/RPE-expression status	74
Figure 3.2	Proportion and classification of genes represented by chromosome 6q12-q16 EST clones hybridizing to human retina/RPE-derived cDNA probes	76
Figure 3.3	Characterization of various chromosome 6q12-q16 retina/RPE-expressed EST clones by northern blot analysis	89
Figure 3.4	Assessment of interspecies sequence conservation of individual retina/RPE-expressed EST clones by Southern blot analysis	91
Figure 3.5	Bioinformatic assessment of evolutionary conservation of retinal/RPE-expressed ESTs mapping to chromosome 6q12-q16	93

Figure 4.1	Cloning strategy used to identify and define the human *C6orf37* gene	110
Figure 4.2	Nucleotide and deduced translation product sequences of human C6orf37 cDNA	112-113
Figure 4.3	Sequence alignment of the human C6orf37 protein and homologues identified in other species	115-116
Figure 4.4	*In vitro* translation product of *C6orf37* coding sequence	118
Figure 4.5	Expression profile of the *C6orf37* gene in human tissues	120
Figure 5.1	Strategy used to identify and define the human *ELOVL4* gene	144
Figure 5.2	Comparison of human ELOVL4 and mouse Elovl4 protein sequences with the product of the yeast *ELO3* gene, Elo3p	146
Figure 5.3	Hydropathy profile analysis of the human ELOVL4 and yeast Elo3p protein sequences	148
Figure 5.4	Expression profile of the *ELOVL4* gene in human tissues	150
Figure 5.5	Segregation of the 790-794delAACTT mutation in families with the STGD3 and adMD phenotypes	154
Figure 6.1	Genomic sequence conservation of the *ELOVL4* gene in diverse species	177
Figure 6.2	Ocular *ELOVL4* gene expression in mammalian species	179
Figure 6.3	Sequence alignment of the human ELOVL4 protein with known orthologues and most similar homologues identified in other species	181-182
Figure 6.4	Phylogenetic tree of ELO protein family members	184
Figure 6.5	Conservation of ELOVL4 protein synthesis in mammalian retinas	186
Figure 6.6	Conservation of sub-retinal ELOVL4 protein distribution	188
Figure 7.1	A schematic diagram depicting the putative functional protein domains in human ELOVL4 as predicted from the amino acid sequence	210
Figure 7.2	Sequence alterations in the *ELOVL4* coding sequence associated with human macular disease introduced into wild-type ELOVL4 cDNA by site-directed mutagenesis	212

Figure 7.3 Sequence alterations affecting putative functional domains in the ELOVL4 protein introduced into wild-type ELOVL4 cDNA by site-directed mutagenesis 214

Figure 7.4 ELOVL4 expression constructs for mammalian cell transfections 216

Figure 7.5 Western blot analysis of GFP-ELOVL4 fusion proteins expressed in COS-7 cells transfected with pEGFP-ELOVL4 constructs 218

Figure 7.6 Immunofluorescence images of COS-7 cells transfected with pEGFP-ELOVL4 cDNA constructs encoding wild-type and mutant ELOVL4 fusion proteins 220

LIST OF ABBREVIATIONS

ABC	ATP-binding cassette
ABCR	retina-specific ATP-binding cassette transport protein
adMD	autosomal dominant atrophic macular dystrophy
AMD/ARMD	age-related macular degeneration
ARE	AU-rich elements
ATP	adenosine triphosphate
BAC	bacterial artificial chromosome
BCAMD	benign concentric annular macular dystrophy
BCKDH	branched chain α-ketoacid dehydrogenase
BLAST	basic local alignment search tool
bp	base pair
BSA	bovine serum albumin
C6orf7	chromosome 6 open reading frame 7
C6orf37	chromosome 6 open reading frame 37
cDNA	complementary DNA
cGMP	cyclic guanosine monophosphate
cM	centimorgan
CMV	cytomegalovirus
CNV	choroidal neovascularization
COP	coat protein
CORD	cone-rod dystrophy
cpm	counts per minute
°C	degrees celcius
dATP	deoxyadenosine-5'-triphosphate
dCTP	deoxycytidine-5'-triphosphate
dd	dominant drusen
del	deletion
dGTP	deoxyguanosine-5'-triphosphate
DAB	diaminobenzidine
DAPI	4'6-diamidino-2-phenylindole-2HCl
DHA	docosahexaenoic acid
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleoside triphosphates
dTTP	deoxythymidine-5'-triphosphate
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EFEMP	EGF-containing fibrillin-like extracellular matrix protein
Elo/ELO	elongase
ELOVL4	elongation of very long-chain fatty acids 4
ER	endoplasmic reticulum
ERG	electroretinogram
EST	expressed sequence tag
GABRR	retinal gamma-aminobutyric acid receptor
GC	guanylate cyclase
GCAP	guanylate cyclase-activating protein
GCL	ganglion cell layer
GFP	green fluorescent protein

GMP	guanosine monophosphate
HCl	hydrochloric acid
IgG	immunoglobulin G
ILM	inner limiting membrane
IMPG	interphotoreceptor matrix proteoglycan
INL	inner nuclear layer
IPL	inner plexiform layer
IPM	interphotoreceptor matrix
IS	inner segment
ISCEV	International Society for the Clinical Electrophysiology of Vision
kb	kilobase
kDa	kiloDalton
LCA	Leber congenital amaurosis
LOD	logarithm of odds
µCi	microcurie
µg	microgram
µl	microlitre
µm	micrometre
M	molar
MCDR	North Carolina macular dystrophy
$MgCl_2$	magnesium chloride
ml	millilitre
mm	millimetre
mM	millimolar
mmol	millimole
mRNA	messenger RNA
NaCl	sodium chloride
NCBI	National Center for Biotechnology Information
ng	nanogram
NH_4Cl	ammonium chloride
NH_4OAc	ammonium acetate
nm	nanometre
OD	oculus dexter (right eye)
OLM	outer limiting membrane
OMIM	Online Mendelian Inheritance in Man
ONL	outer nuclear layer
OPL	outer plexiform layer
ORF	open reading frame
OS	outer segment
PAC	phage P1-derived artificial chromosome
PAGE	polyacrylamide gel electrophoresis
PBCRA	progressive bifocal chorioretinal atrophy
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDE	phosphodiesterase
PDI	protein disulfide isomerase
PEX	peroxin
ph. rec.	photoreceptor
PRL	phosphatase of regenerating liver
RACE	rapid amplification of cDNA ends

RDS	retinal degeneration slow
RIM	Rab3A-interacting molecule
RNA	ribonucleic acid
RP	retinitis pigmentosa
RPE	retinal pigment epithelium
RPE65	retinal pigment epithelium-specific protein, 65 kDa
RPGR	retinitis pigmentosa GTPase regulator
rpm	revolutions per minute
RT-PCR	reverse transcription - polymerase chain reaction
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SSC	sodium chloride, tri-sodium citrate buffer
STGD	Stargardt disease
STS	sequence-tagged site
θ	recombination fraction
TGN	*trans* Golgi network
TIMP	tissue inhibitor of metalloproteinases
Tris	tris(hydroxymethyl)aminomethane
TRITC	tetramethylrhodamine isothiocyanate
TSS	transcription start site
UTR	untranslated region
YAC	yeast artificial chromosome
Zmax	maximum LOD score

Chapter 1

Introduction

Blindness is one of the most feared of human afflictions. It represents one of the major health problems worldwide in terms of human suffering as well as its detrimental impact on work productivity and demands on health care systems. In Canada, over 2.8 million people are affected by some form of visual impairment.[1] Statistics of individuals registered with the Canadian National Institute for the Blind reveal that over one-third of cases are attributable to an established or presumed genetic origin.[2] The eye represents a unique organ in that a high proportion of genetic abnormalities with functional consequences can be detected due to their effects on vision. With the advent of genetic mapping technology and advances in the field of molecular biology, the last two decades have been highly productive in the effort to determine the genetic basis of inherited ocular diseases. Hundreds of disorders with ocular involvement are currently documented in the Online Mendelian Inheritance in Man (OMIM) database that have either disease gene or chromosomal map location assignments.[3] The promise of such progression in understanding disease etiology is the eventual ability to interfere with the defective biological processes that result in vision loss at their most basic, molecular level.

1.1 THE HUMAN EYE

The human eye is a complex organ composed of several different types of tissues that serve distinct functions (Figure 1.1). Its origins lie in two of the three embryological cellular layers: the ectoderm, which gives rise to the nervous system and epithelial cells and hence the major parts of the eye, and the mesoderm, which forms the orbital bones, extraocular muscles, and blood vessels associated with the eye.[4] Each of the fully developed ocular structures must work in coordination and with precise movements, reflexes, and connectivity in order for us to accurately perceive our external environment by sight. The primary series of events in visual perception involve the passage of incident light through the transparent cornea and lens which focus it onto the retina at the back of the eye. The retina converts the light signal into a neuronal impulse, which is then passed through the optic nerve to the lateral geniculate nucleus of the brain, and ultimately reaches the primary visual cortex that is responsible for transforming the information it receives into our sense of vision.[5]

1.2 THE RETINA

The retina corresponds to the transparent neural tissue lining the internal surface of the posterior two-thirds of the eye, located between the vitreous and the vascular choroid layer. It represents an externalized portion of the brain that forms during early embryonic development. The retina plays a fundamental role in vision, serving as the initial receptor for visual stimuli and transmitting the visual information to the brain via the optic nerve. Both functionally and anatomically, the retina is by far the most complex structure in the eye. It is composed of three major classes of cells - neurons, glial cells, and epithelial cells - that are arranged in a highly organized anatomical and hierarchical structure (Figure 1.2). The neurons alone can be subdivided into 50-100 morphologically, physiologically, or histochemically distinguishable cell types,[4] demonstrating the vast cellular heterogeneity and thus functional complexity of this biological tissue.

1.2.1 Retinal Cell Types and Spatial Organization

The major cell types of the mammalian retina are organized into distinct nuclear layers that are separated by intervening layers comprised of neuronal cell processes and synaptic terminals (Figure 1.2). The outermost layer, located in the furthest posterior position relative to the incoming light stimulus, is the retinal pigment epithelium (RPE). RPE cells are not directly involved in the neural events of vision, but are essential for the survival and maintenance of other retinal cell types, namely the photoreceptors. The photoreceptor cell nuclei form the outer nuclear layer (ONL), separated from the RPE by first the outer segment discs followed by the inner segments of the photoreceptors that contain the other cellular organelles. The photoreceptor inner and outer segments extend towards the RPE from the nuclei-containing cell bodies. The photoreceptor axons extend in the opposite direction and comprise the outer plexiform layer (OPL). The inner nuclear layer (INL) consists of the cell bodies of four type of neurons - bipolar cells, horizontal cells, amacrine cells, and interplexiform cells - and also the nuclei of Müller cells, the main glial cells in the retina. Next lies the inner plexiform layer (IPL) that contains the axon terminals of the bipolar cells, neural processes of amacrine and interplexiform cells, and the dendrites of the cells that constitute the following nuclear layer towards the inner retina, the ganglion cells. Within the ganglion cell layer (GCL) are also found some displaced amacrine cells. The next layer inward is the nerve fiber layer, which contains the ganglion cell axons that innervate the optic nerve. Finally, the innermost layer, the inner limiting membrane (ILM), houses the terminal expansions of the Müller cells. The external or outer limiting membrane (OLM) lies along the border between the photoreceptor inner segments and cell bodies of the ONL and corresponds to a series of tight junctions between the photoreceptors and Müller cells.[4]

Each of the neuronal cell types of the retina are involved primarily in either vertical or lateral pathways of nerve impulse transmission, characterizing the direction of information flow between the different cells. The vertical pathways involve the photoreceptors, bipolar cells, and ganglion cells, while horizontal and amacrine cells form the lateral pathways. Interplexiform cells appear to play a role in both vertical and lateral transmission of signals.[4]

1.2.1.1 Photoreceptors

The specialized cell types within the retina responsible for initiating the process of visual transduction and the neural response to light are the photoreceptors. Two types of photoreceptor cells exist, termed rods and cones. The adult human retina contains approximately 92 million rod cells and roughly 5 million cone cells that are distributed non-uniformly.[6] Rods are responsible for vision at low luminescence levels and function in night vision and peripheral vision. Cones are important for vision at high luminescence levels and are required for central vision, visual acuity, and colour perception. Accordingly, the rod photoreceptor cells are concentrated at the peripheral areas of the retina, while cone cells are found in greatest abundance at the central regions.

The cone-rich central area of the retina is referred to as the macula, and the centermost region of the macula is called the fovea (Figure 1.3). The human macula is roughly 5.5 mm in diameter and is histologically distinct from the rest of the retina, containing a higher density of cone photoreceptor cells, ganglion cells and RPE cells than the peripheral retina.[4] Within the macular region there is extreme variation in the number and thickness of the neuronal cell layers (Figure 1.3). The fovea is approximately 1.5 mm in diameter, which accounts for less than 5% of the total area of the retina, and contains purely cone photoreceptors. This region functions as the center for highest visual acuity and spatial resolution; it is the position of the retina to which the observer brings the image of greatest interest in the visual field upon voluntary movement of the eyeball.[7] The foveola or foveal pit forms the central floor of the fovea and has a diameter of 0.35 mm. This is the region of highest concentration of cones in the retina.[7] The foveolar umbo corresponds to a tiny depression in the geographical center of the fovea and is comprised of a single layer of cone cells. The area surrounding the foveola is termed the parafoveal region.

1.2.1.2 Retinal Pigment Epithelium

The retinal pigment epithelium is an amitotic cellular monolayer consisting of melanin-containing epithelial cells that separates the neural retina from the vascular-rich choroid layer of the eye (Figure 1.4). The human retina contains approximately 5 million RPE cells,[8] and like the photoreceptors, cells of the RPE are arranged in a non-uniform distribution across the retina.[9] Although not directly involved in the neural processes underlying vision, the RPE serves a number of critical roles that are essential for proper visual function and photoreceptor cell survival. These include maintenance of the blood-retina barrier, phagocytosis and digestion of photoreceptor cell outer segments, regeneration of photopigment, transport of molecules between the retina and choroid, absorption of excess light to improve image resolution, elimination of metabolic by-products of the visual signal transduction pathway, and maintenance of retinal adhesion.[10,11] Therefore, alterations in the ability of the RPE to perform any of these functions may be involved in causing or contributing to the pathogenesis of retinal diseases.

1.2.2 Phototransduction

When light enters the eye and strikes the retina, a cascade of biochemical reactions occurs within the photoreceptor cells to convert the signal carried by photons into an electrical impulse that can be passed onto the brain via retinal and cerebral neural connections. The signal conversion process is known as phototransduction (Figure 1.5).[12] Light is first absorbed by photopigment molecules located in the outer segment discs of photoreceptor cells, rhodopsin in rod cells and one of three cone opsins (red, green, or blue cone opsin, depending on the wavelength of light maximally absorbed) in cone cells. The photopigment molecule consists of an opsin protein moiety and a covalently attached chromophore, retinal. The photopigment molecule attains its active form upon the light-induced isomerization of the chromophore that results in the conversion of 11-*cis*-retinal to all-*trans*-retinal. The activated photopigment is then able to interact with other molecules in the disc membrane. Interaction with the trimeric G-protein transducin results in its activation, which in turn activates the tetrameric enzyme cyclic guanosine monophosphate (cGMP)-phosphodiesterase (PDE). PDE is then able to catalyze the conversion of cyclic GMP to 5'-GMP, effectively decreasing the intracellular cGMP concentration and subsequently closing the cGMP-gated cation channels in the plasma membranes of photoreceptors. This causes a hyperpolarization, making the membrane potential more negative, and reduces the rate of release of the neurotransmitter, glutamate. This chemical signal then influences the membrane potential of downstream neurons with which the photoreceptors synapse, by either opening or closing ionic membrane channels, and the neuronal signaling cascade is thus initiated.

Inactivation of the phototransduction cascade occurs through termination of the interaction between the activated photopigment and transducin and by dissociation of the all-*trans*-retinal from the opsin molecule. Both processes are mediated by the protein arrestin, which is able to physically bind to the photopigment, and ultimately leads to a decrease in PDE activity, restoration of cellular cGMP levels, and re-opening of the cGMP-gated membrane channels. Regeneration of the 11-*cis*-retinal occurs in the RPE, after which the phototransduction cascade can once again begin upon reassociation of the chromophore with opsin and light exposure.

1.3 INHERITED DISEASES OF THE RETINA

A variety of defects in vision can result from abnormalities in photoreceptor or RPE cell function. Several genetic eye disorders that specifically affect the retina have been described and are collectively referred to as inherited retinopathies. Such disorders represent the most common cause of vision loss leading to irreversible blindness and for which there is at present no known cure. Characterizing these diseases is the phenomenon of retinal degeneration, involving loss of photoreceptor cells through the process of programmed cell death, or apoptosis.[14] The mechanisms underlying the stimulation of apoptosis in retinal cells that results in retinopathy are not fully understood and are currently the focus of active research.

Hereditary retinal degenerative diseases exhibit significant heterogeneity both clinically and genetically.[15] Individual types of photoreceptor degeneration are distinguished by their pattern of cellular and visual loss and corresponding appearance of the retina, mode of inheritance, and, in an increasing number of cases, by the gene defect underlying the disorder.

1.3.1 Diagnosis of Retinal Degenerative Phenotypes

A variety of clinical procedures are employed by ophthalmologists in the diagnosis and phenotypic characterization of retinal diseases. The clinical tests used allow monitoring of both appearance and function of the retina, and can be performed repeatedly over time to track the progressive nature of retinal degenerative disorders.

1.3.1.1 Psychophysical Parameters

The measurement of visual performance often involves the use of tests that rely on verbal responses given by patients upon presentation of a visual stimulus. This type of clinical assessment is used for tests of visual acuity, colour vision, and the visual field.

Visual acuity is measured by determining the resolution threshold of an object that can be viewed by an individual at a given distance. The most familiar acuity test utilizes the Snellen eye chart, consisting of rows of letters of the alphabet that are successively reduced in size. The acuity measurement is reported as the Snellen ratio, where the numerator denotes the distance between the observer and the eye chart, and the denominator specifies the distance from the chart at which a normal individual can distinguish the letters.[16] A ratio of 20/20 is defined as normal visual acuity, indicating that at a test distance of 20 feet, judged to be a practical distance for assessing vision, the patient can reliably identify letters of the chart whose height subtend a 5-minute angle to the eye at 20 feet.[4] If at the same test distance the observer can only view the object if it is twice as large, this indicates that the individual can identify a letter that subtends the same visual angle at 40 feet, and their visual acuity is 20/40. Legal blindness is defined as 20/200 acuity, meaning that an object viewed by such an individual must be 10 times larger in linear dimension than that resolved by someone with 20/20 or normal vision. An alternative description is that a legally blind individual can view an object at a distance of 20 feet that a person with normal visual acuity can see at 200 feet.[16] Visual acuity levels can also be specified by Snell-Sterling's efficiency ratings, which are given as a percentage value.[17] A Snellen ratio of 20/20 is equivalent to 100% visual efficiency, while 20/200 corresponds to a Snell-Sterling rating of 20%.

The most common test of colour vision is the Ishihara pseudo-isochromatic colour plate exam. This is a series of test plates on which printed dots of different colours are used to produce patterns that can be identified as basic objects such as numbers, letters, or shapes.[18] The patterns are created using chromaticities that are confused by colour-deficient individuals, thereby preventing the identification of the objects if the patient is unable to distinguish between different colours. Abnormalities in colour perception are reported as the number of errors made in judging the identity of the object represented on each plate. At least two incorrect plates signify defective colour vision.[18]

1.3.1.2 Fundus Imaging

Fundus imaging, or funduscopy, is a procedure that enables visualization of the back of the eye, termed the fundus, through a low-magnification device known as an ophthalmoscope. Using this technique, the appearance of the retina and its vasculature can be monitored and gross abnormalities such as defects in the RPE, retinal and sub-retinal deposits, neovascularization, and retinal detachments can be detected.[4] The structures visible in a fundus photograph include the optic nerve head, appearing as a yellow disc on either the right or left side depending on which eye is observed (Figure 1.6A). Branches of retinal arteries and veins diverge from the optic disc and arc towards and around the fovea. The macular and foveal regions are also denoted by the darkened area that arises from the presence of xanthophyll pigments in the macula (Figure 1.6A).[4] The peripheral retina surrounds the pigmented central region and is generally a uniform orange-red colour. Degeneration of retinal cells can be detected by the appearance of fundus discolouration such as the loss of pigmentation in the macular region (Figure 1.6B) and the non-uniformity of the surrounding peripheral retina (Figure 1.6C).

1.3.1.3 Fluorescein Angiography

To monitor intraocular blood flow, patients are intravenously injected with a fluorescent dye that reaches the retinal vasculature and can be visualized when the retina is illuminated with ultraviolet light. Fundus photography of the patient produces an image known as a fluorescein angiogram. This technique enables detection of abnormalities in the retinal circulation and blood vessel density. For example, the presence of capillary leakage would result in brighter fluorescence, or hyperfluorescence, compared to that observed for a normal individual, while hypofluorescence may indicate arterial blockage.[4] Neovascularization can also readily be detected with this procedure.

Fluorescein angiography is also useful for identifying RPE defects, since the fluorescent signals generated by the choroidal vasculature are normally attenuated by the melanin within RPE cells. Thus the abnormal appearance of fluorescence in regions that normally would appear dark, such as the fovea (due to the absence of retinal capillaries, the presence of macular pigment, and a greater density of RPE cells in this area), may suggest RPE cell atrophy or loss in those regions.[19] In contrast, RPE abnormalities can also occlude fluorescent signals emanating from retinal blood vessels. In some conditions, visibility of the choroidal circulation is largely abolished. This phenomenon is referred to as a "dark" or "silent" choroid, and is characteristic of autosomal recessive Stargardt disease.[20] In such cases, it is believed that the choroidal fluorescence is blocked by the pathologic accumulation of lipofuscin or lipofuscin-like material, thought to represent cellular debris and waste products, within RPE cells.[20]

1.3.1.4 Electroretinography

The electrical response generated by the retina upon exposure to light can be measured by placing an electrode on the corneal surface and a reference electrode on the skin. The signal recorded as the voltage difference upon illumination is referred to as the electroretinogram (ERG) and consists of sequential waves of electrical activity that originate in the different layers and cell types of the retina.[4] By comparing the waveforms generated from a patient's eye to those recorded for a normal individual, dysfunction of particular retinal cell types in the transmission of the visual signal through the retina can be assessed. The a-wave is the first waveform appearing immediately after presentation of the light stimulus and corresponds to the activity of photoreceptor cells. The b-wave is believed to reflect the neuronal impulses occurring within the inner retina, postsynaptic to the photoreceptors, and is attributed predominantly to bipolar cells. The c-wave represents a slow voltage change that originates in the RPE. Although RPE cells are not neurons, an electrical potential is generated across this neuroepithelial layer as a result of ion fluxes around the closely associated photoreceptor cells.

The ERG can also selectively measure rod or cone responses. Under scotopic or dark-adapted conditions, the response is largely mediated by rod cells, whereas cone-dominated responses are obtained under photopic or light-adapted conditions.[4]

1.3.2 Regional Variation of Retinal Degeneration

Preferential degeneration of either of the two photoreceptor cell types specifies either peripheral or central retinopathies. The most frequent inherited forms of visual handicap are represented by the collection of retinopathies known as retinitis pigmentosa that target the peripheral retina. These disorders primarily affect rod function and as a result lead to night blindness and narrowing of the visual field ("tunnel vision") due to peripheral vision defects. Macular dystrophies are a heterogeneous group of retinal diseases primarily affecting the central cone-rich foveal-macular region and are largely responsible for the prevalence of legal blindness in elderly populations in industrialized countries.[21] Macular degeneration results in decreased visual acuity, impairments in colour perception, and overall reduction in central vision.

In addition to the nature of the vision loss experienced, clinical examination of retinal dystrophy patients involving other diagnostic procedures can distinguish between cases of peripheral and central degeneration. Selective cell deterioration in the peripheral retina or the macular region can easily be distinguished by fundus imaging (Figure 1.6B and 1.6C). Fluorescein angiography can also reveal macular and foveal defects resulting from RPE cell loss. Multifocal electroretinography can further assess regional functional differences across the retina.[4]

1.3.3 Genetic Diversity of Retinal Degeneration

The genetic diversity of retinal degenerative disorders is apparent from their autosomal dominant, autosomal recessive, X-linked, and mitochondrial inheritance patterns observed. Additionally, different mutations in distinct genes have been identified as the causative factors for a growing number of cases of inherited retinal degeneration. The retinitis pigmentosa category of diseases is the most genetically well-characterized, with mutations in 24 different genes and several other uncharacterized disease loci identified to date.[22] The nature of the responsible genetic defects in some forms of macular degeneration have also been elucidated, with mutations in 7 disease genes reported thus far.[22,23] An additional 72 genes are currently known to underlie other retinal degenerative phenotypes including cone-rod dystrophies, congenital stationary night blindness, Leber congenital amaurosis, and Usher syndrome.[22]

As might be predicted, the majority of retinal disease genes identified encode retina-specific proteins involved in the visual process, such as components of the phototransduction cascade. However, mutations in a number of widely expressed genes have more recently been found to cause retinal degenerative conditions,[24] for which the mechanisms of selective retinal disease pathogenesis are not as readily apparent. While the advances in elucidating the genetic basis of retinal dystrophies has been rapid, there still remain many clinically described inherited retinal diseases for which the cause is yet unknown.

1.3.4 Age-Related Macular Degeneration

Age-related macular degeneration (AMD) is the leading cause of blindness worldwide in individuals over the age of 60, with an estimated 25-30 million people afflicted with this disease.[25-27] AMD is a progressive form of retinal degeneration that primarily affects the macular region, and encompasses a broad spectrum of clinical findings.[28] Two distinct forms of AMD have been designated based on ocular features at late stages of the disease: dry AMD (also known as non-exudative AMD, geographic atrophy, or atrophic AMD) and wet AMD (also referred to as exudative, disciform, or neovascular AMD, or choroidal neovascularization (CNV)). The dry form of AMD is characterized by the gradual loss of RPE and photoreceptor cells and resultant retinal dysfunction experienced as slow visual deterioration over time. Wet AMD involves the aberrant formation of new blood vessels in the choroid layer that leads to haemorrhaging, detachment of the retina from the RPE, and fibrotic scarring in the macular region. In this form of the disease, vision defects appear more suddenly and the loss of vision occurs more rapidly. Dry AMD is more prevalent, occurring in roughly 80-90% of AMD patients, while the remaining 10-20% have the wet form.[29] Although much less common, wet AMD is responsible for the majority of cases of severe vision loss, causing nearly 90% of legal blindness attributable to AMD.[19] A number of treatments for AMD have been developed for the wet form of the disease that target the choroidal neovascularization that occurs.[30] However these treatments are not appropriate for all wet AMD patients, and currently there are no treatments available for dry AMD patients, demonstrating the paucity of therapeutic options for most individuals suffering from AMD.

AMD is a complex, multifactorial disorder for which the cause is unknown, but is thought to arise from the interaction of genetic and environmental factors. The strongest risk factors are age and family history, and lifestyle factors such as diet, smoking, alcohol consumption, and light exposure may also influence manifestation of the disease.[31,32] Family and twin studies suggest the likelihood of a genetic component contributing to AMD pathogenesis,[33-37] and a number of linkage studies have attempted to map the location of AMD susceptibility genes.[38-46] A recent report describes the identification of a sequence variation in the *HEMICENTIN-1* gene segregating with AMD in a family for which the ARMD1 locus was defined.[47] However it is unclear if this is a true disease-causing mutation or merely a marker for the disease haplotype.[46] The late onset of the disease phenotype, complex genetics, and multifactorial nature of the disorder present difficulties for performing genetic mapping studies of AMD. Through the investigation of the more rare, earlier-onset, monogenic, but clinically similar forms of inherited macular dystrophy, insight may be gained into the molecular pathogenesis and mechanisms of progression of maculopathy in cases of AMD.

1.4 FOCUS OF THE THESIS

The work that is documented in this thesis was initiated with the aim of contributing to the state of knowledge on retinal degenerative diseases and their molecular foundations. In Chapter 2 a Canadian family is described in which a form of macular degeneration had been identified and was followed over several generations. Genetic analysis of the family revealed linkage of the disease phenotype to a region of human chromosome 6 that was known to harbour multiple disease loci, thereby defining a genomic retinal dystrophy "hotspot". Towards the continued search for the disease gene in the Canadian macular dystrophy pedigree, a candidate gene approach was adopted to identify all the retinal or retinal pigment epithelial genes residing in the chromosome 6q interval defined in the positional cloning studies. Establishment of shared ancestry between the Canadian family and a kindred based in the United States with a similar clinical form of macular disease enabled narrowing of the genomic interval of interest and thus further reduced the number of candidate genes to assess. Chapter 3 describes the development of a retinal and RPE transcript map for the 6q12-q16 interval that served as a resource for defining retinal disease-relevant genes. The identification and characterization of one such gene, *chromosome 6 open reading frame 37* (*C6orf37)*, is discussed in Chapter 4. This gene proved not to be the defective gene in affected members of the Canadian family, however its discovery adds to current knowledge of retinal molecular biology. Chapter 5 outlines the procedures that led to the ultimate isolation of the disease gene in question, *elongation of very long-chain fatty acids 4* (*ELOVL4)*. The process once again involved in part a collaborative effort in which several independently ascertained pedigrees were found to be related and from which the genetic data was pooled to further pinpoint the chromosomal location of the causal gene. Extensive characterization of the disease gene using both bioinformatic and traditional experimental methods is presented in Chapters 5, 6, and 7, yielding novel and valuable information on the newly defined gene whose function is critical for retinal cell integrity and visual processes. Finally, an initial examination of possible pathogenic mechanisms in patients harbouring *ELOVL4* mutations is addressed in Chapter 7. Altered cellular characteristics of mutant forms of the encoded protein were observed. These studies provide the first clues as to the molecular events that ultimately lead to a retinal degenerative phenotype in the patients from the Canadian family and their now known relatives.

This thesis therefore documents the course of our investigation that began in the clinic and now continues in the laboratory towards eventually coming full-circle with the potential of therapeutic intervention of disease progression based on our experimental findings.

1.5 REFERENCES

1. Statistics Canada 1991 census results. World Wide Web URL: http://www.accessdome.com/com-general/can-stats.asp .

2. Johnson GJ, Green JS. (1987) Prevention of blindness due to genetic disease. *Can. J. Ophthalmol.* 22:145-151.

3. Online Mendelian Inheritance in Man, OMIM (TM). McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD), 2000. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/ .

4. Oyster CW. (1999) *The Human Eye: Structure and Function.* Sunderland, MA: Sinauer Associates. 766 pp.

5. Hubel DH. (1995) *Eye, Brain, and Vision.* New York, NY: Scientific American Library. pp.59-71.

6. Curcio CA, Sloan KR, Kalina RE, Hendrickson AE. (1990) Human photoreceptor topography. *J. Comp. Neurol.* 292:497-523.

7. Cohen AI. (1992) The Retina. In Hart WM Jr (ed): *Adler's Physiology of the Eye: Clinical Application.* St. Louis: Mosby-Year Book. p.594.

8. Gordon W, Bazan N. (1997) Retina. In Harding J (ed): *Biochemistry of the Eye.* London: Chapman & Hall. pp.144-275.

9. Snodderly DM, Sandstrom MM, Leung IY, Zucker CL, Neuringer M. (2002) Retinal pigment epithelial cell distribution in central retina of rhesus monkeys. *Invest. Ophthalmol. Vis. Sci.* 43:2815-2818.

10. Marmorstein AD, Finnemann SC, Bonilha VL, Rodriguez-Boulan E. (1998) Morphogenesis of the retinal pigment epithelium: toward understanding retinal degenerative diseases. *Ann. N. Y. Acad. Sci.* 857:1-12.

11. Boulton M, Dayhaw-Barker P. (2001) The role of the retinal pigment epithelium: topographical variation and ageing changes. *Eye* 15:384-389.

12. Ridge KD, Abdulaev NG, Sousa M, Palczewski K. (2003) Phototransduction: crystal clear. *Trends Biochem. Sci.* 28:479-487.

13. Polans A, Baehr W, Palczewski K. (1996) Turned on by Ca^{2+}! The physiology and pathology of Ca^{2+}-binding proteins in the retina. *Trends Neurosci.* 19:547-554.

14. Travis GH. (1998) Mechanisms of cell death in the inherited retinal degenerations. *Am. J. Hum. Genet.* 62:503-508.

15. Sullivan LS, Daiger SP. (1996) Inherited retinal degeneration: exceptional genetic and clinical heterogeneity. *Mol. Med. Today* 2:380-386.

16. Stein HA, Slatt BJ, Stein RM. (1992) *A Primer in Ophthalmology.* St. Louis: Mosby-Year Book. pp.56-57.

17. Westheimer G. (1992) Visual Acuity. In Hart WM Jr (ed): *Adler's Physiology of the Eye: Clinical Application.* St. Louis: Mosby-Year Book. pp.539-540.

18. Parr J. (1989) *Introduction to Ophthalmology*, 3rd ed. Oxford: Oxford University Press. p.73.

19. Kanski JJ, Milewski SA. (2002) *Diseases of the Macula: A Practical Approach*. London: Harcourt. pp.9,19.

20. Eagle RC Jr, Lucier AC, Bernardino VB Jr, Yanoff M. (1980) Retinal pigment epithelial abnormalities in fundus flavimaculatus: a light and electron microscopic study. *Ophthalmology* 87:1189-1200.

21. van Leeuwen R, Klaver CC, Vingerling JR, Hofman A, de Jong PT. (2003) Epidemiology of age-related maculopathy: a review. *Eur. J. Epidemiol.* 18:845-854.

22. Daiger SP. RetNet, the Retinal Information Network http://www.sph.uth.tmc.edu/Retnet/ .

23. Michaelides M, Hunt DM, Moore AT. (2003) The genetics of inherited macular dystrophies. *J. Med. Genet.* 40:641-650.

24. Pacione LR, Szego MJ, Ikeda S, Nishina PM, McInnes RR. (2003) Progress toward understanding the genetic and biochemical mechanisms of inherited photoreceptor degenerations. *Annu. Rev. Neurosci.* 26:657-700.

25. Klein R, Klein BE, Linton KL. (1992) Prevalence of age-related maculopathy. The Beaver Dam Eye Study. *Ophthalmology* 99:933-943.

26. Vingerling JR, Dielemans I, Hofman A, Grobbee DE, Hijmering M, Kramer CF, de Jong PT. (1995) The prevalence of age-related maculopathy in the Rotterdam Study. *Ophthalmology* 102:205-210.

27. Mitchell P, Smith W, Attebo K, Wang JJ. (1995) Prevalence of age-related maculopathy in Australia. The Blue Mountains Eye Study. *Ophthalmology* 102:1450-1460.

28. Ambati J, Ambati BK, Yoo SH, Ianchulev S, Adamis AP. (2003) Age-related macular degeneration: etiology, pathogenesis, and therapeutic strategies. *Surv. Ophthalmol.* 48:257-293.

29. Silvestri G. (1997) Age-related macular degeneration: genetics and implications for detection and treatment. *Mol. Med. Today* 3:84-91.

30. Hooper CY, Guymer RH. (2003) New treatments in age-related macular degeneration. *Clin. Experiment. Ophthalmol.* 31:376-391.

31. Evans JR. (2001) Risk factors for age-related macular degeneration. *Prog. Retin. Eye Res.* 20:227-253.

32. Smith W, Assink J, Klein R, Mitchell P, Klaver CC, Klein BE, Hofman A, Jensen S, Wang JJ, de Jong PT. (2001) Risk factors for age-related macular degeneration: Pooled findings from three continents. *Ophthalmology* 108:697-704.

33. Klein ML, Mauldin WM, Stoumbos VD. (1994) Heredity and age-related macular degeneration: Observations in monozygotic twins. *Arch. Ophthalmol.* 112:932-937.

34. Meyers SM, Greene T, Gutman FA. (1995) A twin study of age-related macular degeneration. *Am. J. Ophthalmol.* 120:757-766.

35. Seddon JM, Ajani UA, Mitchell BD. (1997) Familial aggregation of age-related maculopathy. *Am. J. Ophthalmol.* 123:199-206.

36. Keverline MR, Mah TS, Keverline PO, Gorin MB. (1998) A practice-based survey of familial age-related maculopathy. *Ophthalmic Genet.* 19:19-26

37. Hammond CJ, Webster AR, Snieder H, Bird AC, Gilbert CE, Spector TD (2002) Genetic influence on early age-related maculopathy: a twin study. *Ophthalmology* 109:730-736

38. Klein ML, Schultz DW, Edwards A, Matise TC, Rust K, Berselli CB, Trzupek K, Weleber RG, Ott J, Wirtz MK, Acott TS. (2000) Age-related macular degeneration: Clinical features in a large family and linkage to chromosome 1q. *Arch. Ophthalmol.* 116:1082-1088.

39. Weeks DE, Conley YP, Mah TS, Paul TO, Morse L, Ngo-Chang J, Dailey JP, Ferrell RE, Gorin MB. (2000) A full genome scan for age-related maculopathy. *Hum. Mol. Genet.* 9:1329-1349.

40. Weeks DE, Conley YP, Tsai HJ, Mah TS, Rosenfeld PJ, Paul TO, Eller AW, Morse LS, Dailey JP, Ferrell RE, Gorin MB. (2001) Age-related maculopathy: an expanded genome-wide scan with evidence of susceptibility loci within the 1q31 and 17q25 regions. *Am. J. Ophthalmol.* 132:682-692.

41. Majewski J, Schultz DW, Weleber RG, Schain MB, Edwards AO, Matise TC, Acott TS, Ott J, Klein ML. (2003) Age-related macular degeneration--a genome scan in extended families. *Am. J. Hum. Genet.* 73:540-550.

42. Schick JH, Iyengar SK, Klein BE, Klein R, Reading K, Liptak R, Millard C, Lee KE, Tomany SC, Moore EL, Fijal BA, Elston RC. (2003) A whole-genome screen of a quantitative trait of age-related maculopathy in sibships from the Beaver Dam Eye Study. *Am. J. Hum. Genet.* 72:1412-1424.

43. Seddon JM, Santangelo SL, Book K, Chong S, Cote J. (2003) A genomewide scan for age-related macular degeneration provides evidence for linkage to several chromosomal regions. *Am. J. Hum. Genet.* 73:780-790.

44. Iyengar SK, Song D, Klein BE, Klein R, Schick JH, Humphrey J, Millard C, Liptak R, Russo K, Jun G, Lee KE, Fijal B, Elston RC. (2004) Dissection of genomewide-scan data in extended families reveals a major locus and oligogenic susceptibility for age-related macular degeneration. *Am. J. Hum. Genet.* 74:20-39.

45. Kenealy SJ, Schmidt S, Agarwal A, Postel EA, De La Paz MA, Pericak-Vance MA, Haines JL. (2004) Linkage analysis for age-related macular degeneration supports a gene on chromosome 10q26. *Mol. Vis.* 10:57-61.

46. Abecasis GR, Yashar BM, Zhao Y, Ghiasvand NM, Zareparsi S, Branham KE, Reddick AC, Trager EH, Yoshida S, Bahling J, Filippova E, Elner S, Johnson MW, Vine AK, Sieving PA, Jacobson SG, Richards JE, Swaroop A. (2004) Age-related macular degeneration: a high-resolution genome scan for susceptibility loci in a population enriched for late-stage disease. *Am. J. Hum. Genet.* 74:482-494.

47. Schultz DW, Klein ML, Humpert AJ, Luzier CW, Persun V, Schain M, Mahan A, Runckel C, Cassera M, Vittal V, Doyle TM, Martin TM, Weleber RG, Francis PJ, Acott TS. (2003) Analysis of the *ARMD1* locus: evidence that a mutation in *HEMICENTIN-1* is associated with age-related macular degeneration in a large family. *Hum. Mol. Genet.* 12:3315-3323.

Figure 1.1 Schematic diagram of the human eye. Figure adapted from www.nei.nih.gov/eyediag.htm.

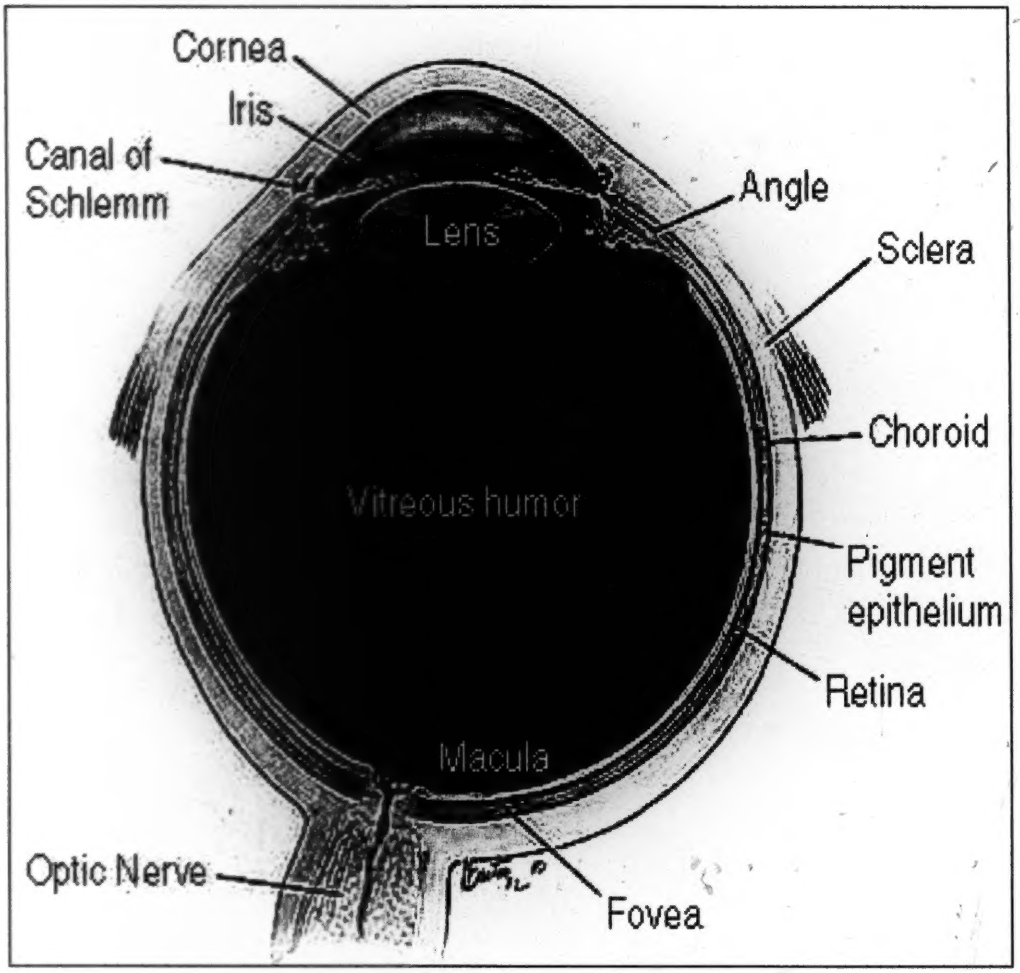
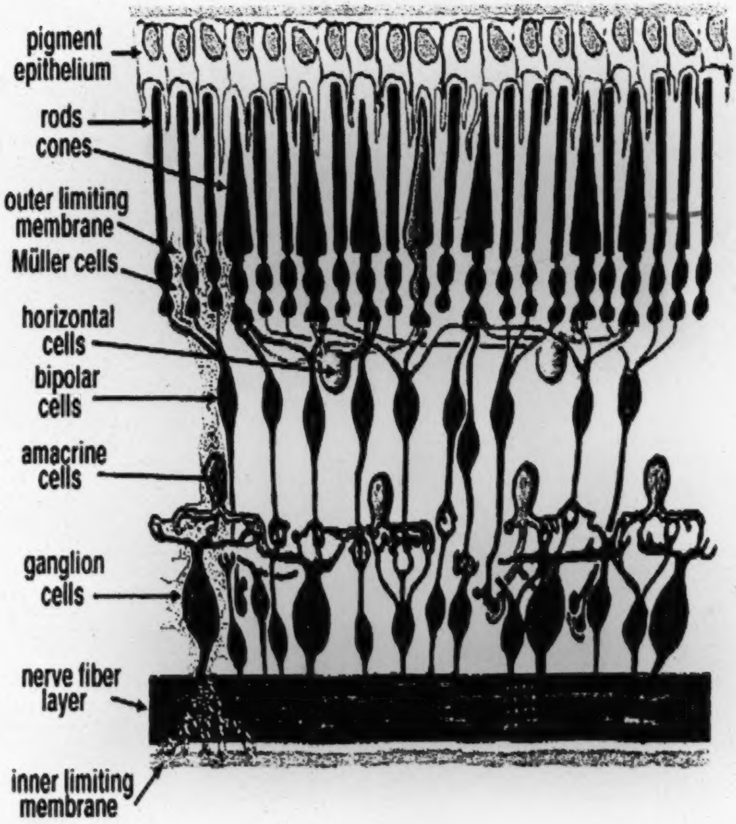

Cornea
Iris
Canal of Schlemm
Angle
Lens
Sclera
Choroid
Vitreous humor
Pigment epithelium
Retina
Macula
Optic Nerve
Fovea

Figure 1.2 Structure of the retina. *A*, Schematic representation depicting the various retinal cell types and their arrangement into organized layers. *B*, Light micrograph of a vertical cross-section through the human retina. *OLM*, outer limiting membrane; *ONL*, outer nuclear layer; *OPL*, outer plexiform layer; *INL*, inner nuclear layer; *IPL*, inner plexiform layer; *GCL*, ganglion cell layer; *ILM*, inner limiting membrane. Images adapted from http://webvision.med.utah.edu.

A


pigment epithelium
rods
cones
outer limiting membrane
Müller cells
horizontal cells
bipolar cells
amacrine cells
ganglion cells
nerve fiber layer
inner limiting membrane

B

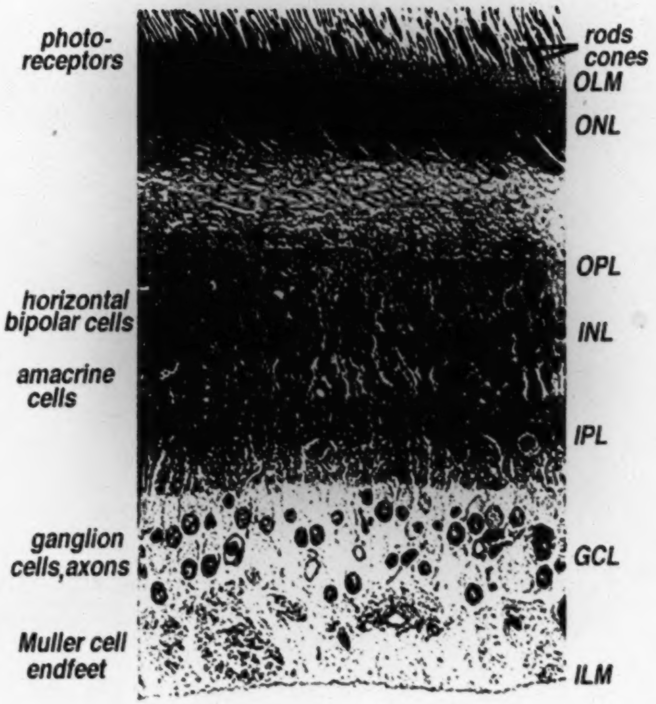


photo-receptors
horizontal bipolar cells
amacrine cells
ganglion cells, axons
Muller cell endfeet
rods
cones
OLM
ONL
OPL
INL
IPL
GCL
ILM

Figure 1.3 Geography of the macular and foveal regions of the human retina. *A*, Light micrograph of a histological cross-section of the human peripheral retina, comprised of a multiple-layered array of neural cells. *B*, Light micrograph of a histological cross-section of the human macular region of the retina. *RPE*, retinal pigment epithelium; *ONL*, outer nuclear layer; *INL*, inner nuclear layer; *M*, outer rim of the macula. Micrographs provided by Dr. Steve L. Bernstein (University of Maryland).

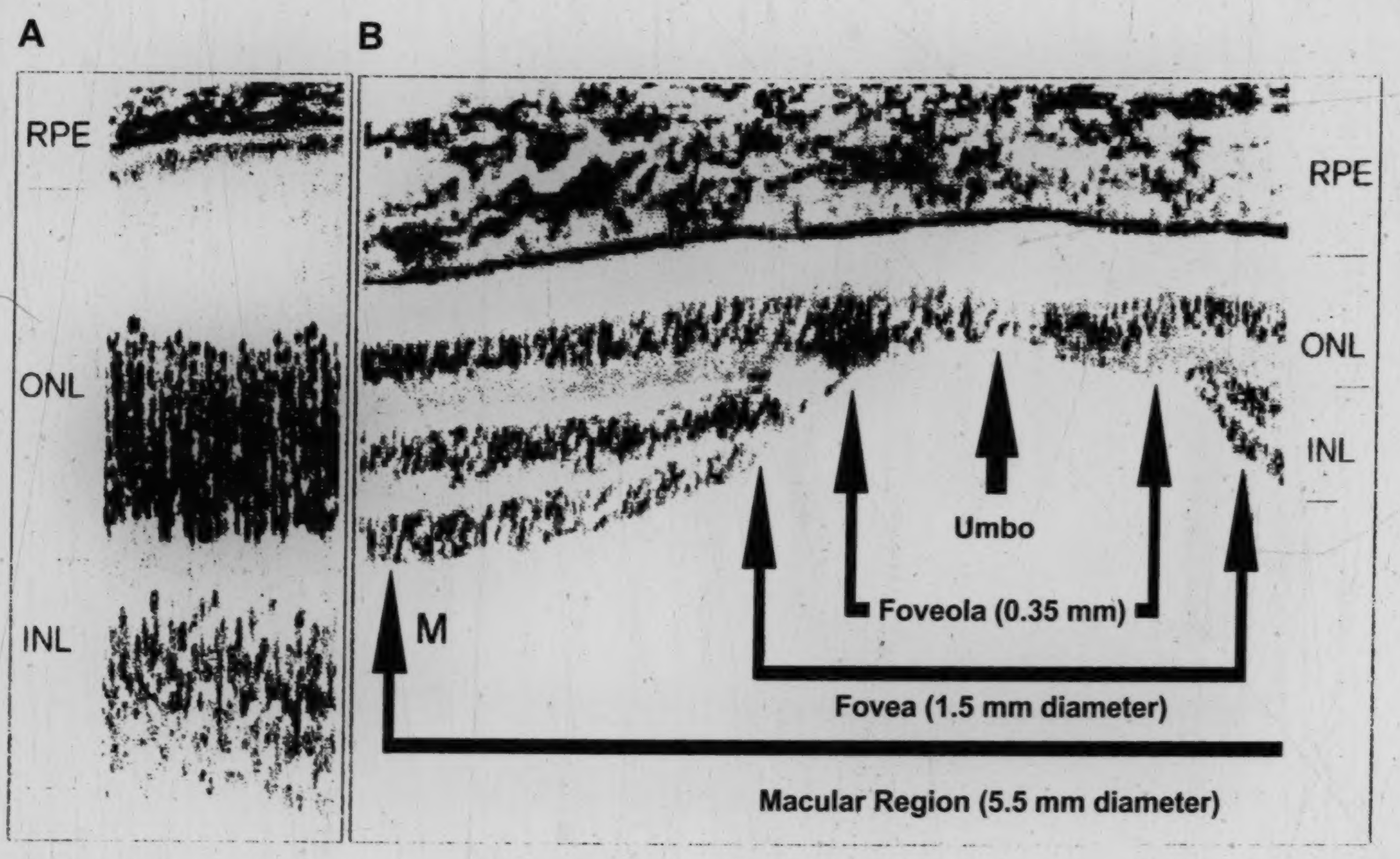

A
B
RPE
ONL
INL
RPE
ONL
INL
Umbo
Foveola (0.35 mm)
Fovea (1.5 mm diameter)
M
Macular Region (5.5 mm diameter)

Figure 1.4 Schematic diagram of the retinal pigment epithelium. The section of the eye highlighted in Panel A is magnified as a light micrograph image in Panel B. Panel C illustrates the position of the photoreceptor cell outer segments relative to the cells of the RPE. *ph. rec.*, photoreceptor. Figure adapted from ref. 10.

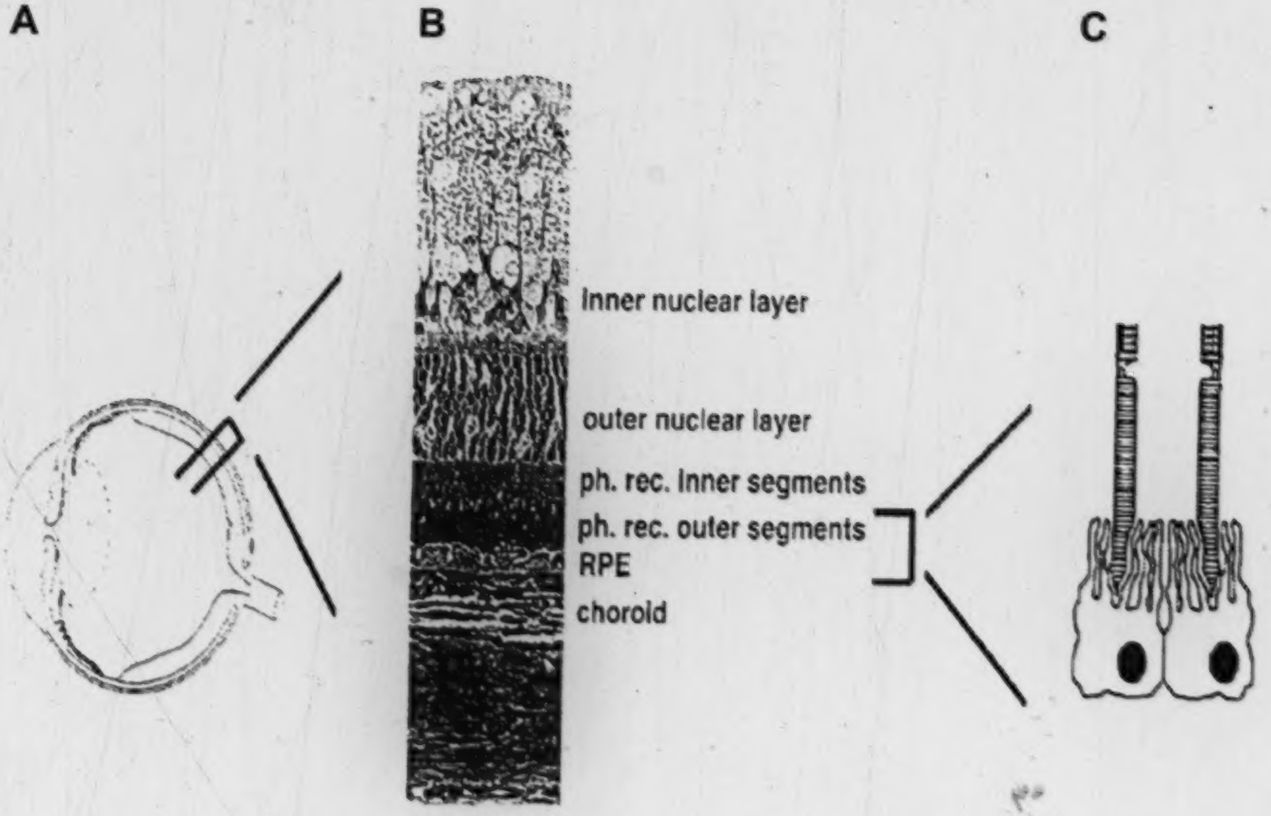

A
B
C
Inner nuclear layer
outer nuclear layer
ph. rec. Inner segments
ph. rec. outer segments
RPE
choroid

Figure 1.5 The vertebrate visual phototransduction cascade. The molecular signal transduction pathway mediating visual input conversion to a chemical signal for neuronal propagation occurring in a typical rod photoreceptor cell is shown. Refer to text for details. *R**, activated rhodopsin; T_α^*, activated transducin; *GDP*, guanosine diphosphate; *GTP*, guanosine triphosphate; *cGMP*, cyclic guanosine monophosphate; *PDE*, cGMP-phosphodiesterase; *PDE**, activated phosphodiesterase; *GC*, guanylate cyclase; *GCAP*, guanylate cyclase-activating protein; *Ca*, calcium; *H+*, proton; *IPM*, interphotoreceptor matrix. Figure adapted from ref. 13.

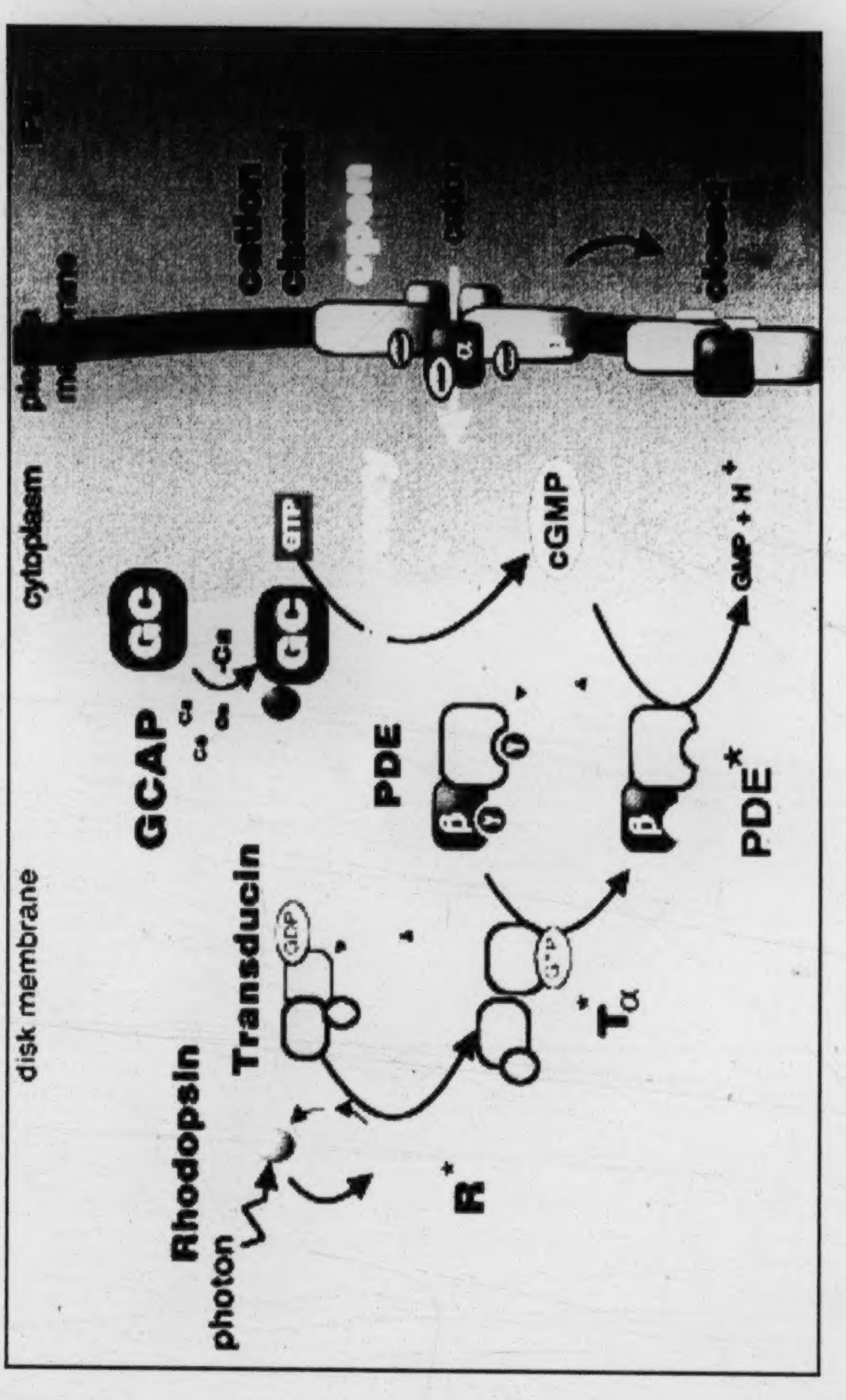
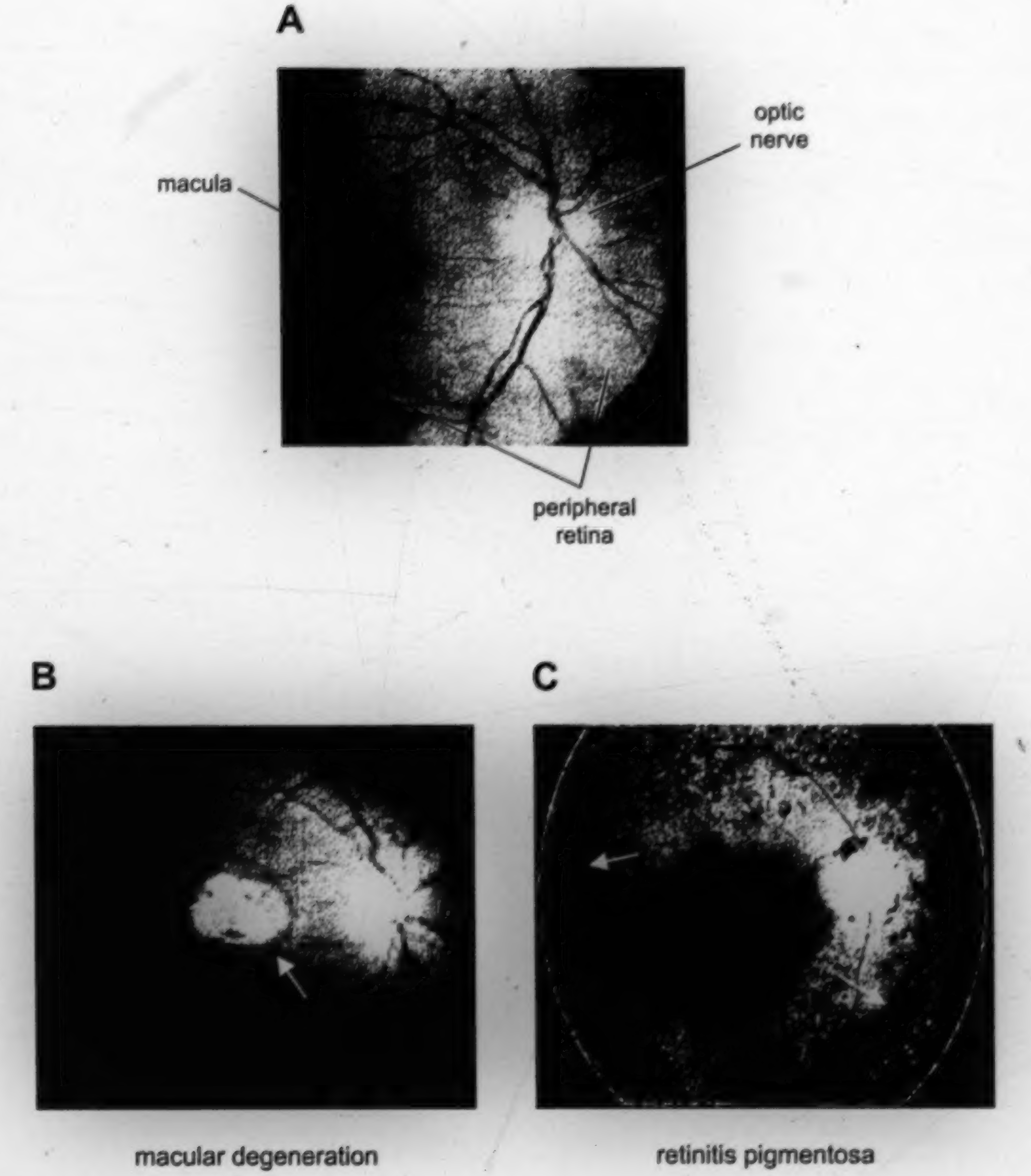

disk membrane
cytoplasm
plasma membrane
cation channel
open
cation
closed
Rhodopsin
photon
R*
Transducin
GDP
GTP
T*α
PDE
β
γ
PDE*
GCAP
Ca
GC
-Ca
GTP
cGMP
GMP + H+
α

Figure 1.6 Funduscopy of the retina. Fundus photographs showing the appearance of the retina in normal individuals (Panel A), in a macular degeneration patient (Panel B), and in a retinitis pigmentosa patient (Panel C). Arrows indicate atrophic regions. Note the regional differences in cell loss distribution between Panels B and C.


A
optic nerve
macula
peripheral retina
B
C

macular degeneration

retinitis pigmentosa

Chapter 2

Autosomal Dominant Stargardt-Like Macular Dystrophy Segregating in a Large Canadian Family

A version of this chapter has been published.
Lagali PS, MacDonald IM, Griesinger IB, Chambers ML, Ayyagari R, Wong PW. Autosomal dominant Stargardt-like macular dystrophy segregating in a large Canadian family. *Canadian Journal of Ophthalmology* 2000;35(6):315-324.

2.1 INTRODUCTION

Macular dystrophies are the leading cause of acquired visual impairment leading to irreversible blindness in the developed world.[1] In Canada alone, over half of the cases of registered blindness are attributed to some form of macular degeneration (K. Boljkovac, CNIB Resource Centre, Toronto, personal communication). Age-related macular degeneration is the most common cause of vision loss in North America in individuals over the age of 50.[2]

Macular degenerative diseases represent a heterogeneous group of disorders characterized by progressive bilateral deterioration of central vision, and atrophy of the macula and underlying retinal pigment epithelium. More than 30 genetically distinct groups of macular dystrophies have been documented.[3] Macular disorders vary in age of onset and clinical presentation, however all share the same outcome: photoreceptor and RPE cell loss in the central retina. Therefore, analysis of a particular form of macular degeneration may contribute to the understanding of the pathogenesis of other forms of macular dystrophy.

Stargardt macular dystrophy is the most common hereditary macular dystrophy, with an incidence of approximately 1 in 10,000 live births and once stated as accounting for 7% of all retinal dystrophies.[4] This disorder is characterized by impaired central vision detected in the first few decades of life, atrophy of the macula and RPE, and the presence of yellow flecks in the macular region resulting from an accumulation of lipofuscin-like material.[5-9] Notably, these clinical features are also observed in cases of AMD.[10] Classical Stargardt dystrophy is inherited as an autosomal recessive disorder, although a number of cases involving autosomal dominant transmission have been described.[9,11-16] The gene for autosomal recessive Stargardt dystrophy has been mapped to chromosome 1p21-p13 (STGD1) and encodes the ABCR protein, a retina-specific ATP-binding cassette transport protein found in rod outer segment disc membranes.[17,18] Three distinct genetic loci for autosomal dominant Stargardt-like disorders have been mapped to date: STGD2 on chromosome 13q34,[13] STGD3 on chromosome 6q14,[12,14] and STGD4 on chromosome 4p.[15] STGD3 is particularly interesting, as the disease locus lies in a region known to contain at least seven other retinal disease loci: cone-rod dystrophy (CORD7),[19] autosomal recessive retinitis pigmentosa (RP25),[20,21] North Carolina macular dystrophy (MCDR1),[22,23] progressive bifocal chorioretinal atrophy (PBCRA),[24] Leber congenital amaurosis (LCA5),[25] dominant drusen with macular degeneration,[26] and another form of autosomal dominant Stargardt-like disease (adMD).[27] This suggests that chromosome 6q represents a "hotspot" for retinal disorders.

The STGD3 locus maps to a 18-centimorgan (cM) interval on chromosome 6q, flanked by microsatellite markers D6S313 and D6S252.[12] This genomic interval overlaps those that have been defined for four other macular diseases, CORD7, LCA5, dominant drusen with macular degeneration, and adMD.[19,25-27] In addition there are three other retinal dystrophy loci which do not overlap the STGD3 locus: RP25, MCDR1, and PBCRA.[20-24] Furthermore, a variety of retinal disorders have been associated with mutations found in the retinal degeneration slow (RDS)/peripherin gene localized to the short arm of chromosome 6.[28-33] Thus it appears that human chromosome 6 contains a number of distinct genes whose dysfunction leads to different types of retinal degenerative disease.

In the current communication, a clinically distinct form of autosomal dominant Stargardt-like dystrophy segregating in a five-generation Canadian family is described. Standard ophthalmologic examination of 15 affected individuals revealed clinical features similar to previously documented forms of Stargardt-like macular degeneration as well as distinguishing characteristics of the disorder affecting this particular kindred. Genetic analysis was subsequently performed using microsatellite markers located on chromosome 6q. Linkage was established between the disease locus and marker locus D6S300 with a two-point maximum LOD score (*Zmax*) of 5.50 at a recombination fraction of 0.00. Therefore, the disease locus in this family maps to a chromosomal region to which other retinal disease loci, including at least two forms of autosomal dominant Stargardt-like macular dystrophy, have also been mapped. Further analysis using a total of 43 chromosome 6q markers defined the disease haplotype.

2.2 MATERIALS AND METHODS

2.2.1 Subjects

Twenty-seven members of a single five-generation, non-consanguineous Caucasian family were identified for study by Dr. Ian M. MacDonald and Dr. William G. Pearce (University of Alberta, Edmonton, AB). Detailed medical and family histories were obtained, and a pedigree was constructed (Figure 2.1). Ophthalmologic examination of family members performed by Dr. I.M. MacDonald and Dr. W.G. Pearce included Snellen visual acuity measurements, Ishihara colour vision testing, and funduscopy. Individuals were designated as affected based on the presence of decreased visual acuity and macular atrophy. Fluorescein angiography was performed on 3 patients, and 6 affected subjects underwent electroretinography and dark adaptometry. Electrophysiological testing was performed on individual V-13 (age 14) according to International Society for the Clinical Electrophysiology of Vision (ISCEV) standards.[34] The b-wave amplitudes of the cone and rod electroretinograms were equally reduced by approximately 40% while the a-wave was only reduced in scotopic conditions.

2.2.2 Genomic DNA Isolation

Blood samples were collected from twenty-four of the twenty-seven individuals examined. White blood cells were isolated from the samples by removing plasma and red blood cells lysed in 0.14 M NH_4Cl, 17 mM Tris-HCl (pH 7.65). Genomic DNA was isolated from the white blood cell fraction by cell lysis in 0.2% SDS, 1 M NaCl, 40 mM EDTA (pH 8.0), 100 mM Tris-HCl (pH 8.0) followed by DNA extraction using a phenol/chloroform-based organic extraction procedure. Genomic DNA was subsequently precipitated from the aqueous phase in 0.4 M NH_4OAc and 95% ethanol and dissolved in 10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA (pH 8.0). DNA was quantitated using a 260 nm wavelength on a Milton Roy Spectronic 601 spectrophotometer (Milton Roy Company, Rochester, NY).

2.2.3 Genotyping

DNA samples were analyzed by polymerase chain reaction (PCR)-based amplification of polymorphic short tandem repeat microsatellite marker sequences located on human chromosome 6 using commercially available primers (Research Genetics, Inc., Huntsville, AL). PCR was performed using a GeneAmp PCR kit (Perkin Elmer, Mississauga, ON) and carried out in 15-microlitre reactions containing 4 ng/μl primers, 0.2 mM dATP, 0.2 mM dGTP, 0.2 mM dTTP, 0.024 mM dCTP, 18 μCi/mmol $[\alpha^{32}P]$-dCTP, 1.5 mM $MgCl_2$, 0.03 units of AmpliTaq DNA polymerase (Perkin Elmer), and 100 nanograms of genomic DNA as template. PCR products were amplified in a

GeneAmp 9600 PCR System Thermocycler (Perkin Elmer) with an initial denaturation of 3 minutes at 94°C followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds, and extension at 72°C for 40 seconds, and a final 3-minute extension at 72°C. Following amplification, PCR products were electrophoresed on a denaturing 6% polyacrylamide gel using a sequencing gel electrophoresis apparatus (International Biotechnology Instruments, New Haven, CT). Bands in the gel corresponding to the PCR products were visualized by exposure of the dried gel to X-ray film (Eastman Kodak, Rochester, NY).

2.2.4 Linkage Analysis

Two-point linkage analysis was performed between the disease locus and each microsatellite marker using the MLINK program (version 5.1) of the LINKAGE software package.[35] Linkage was assessed under conditions of autosomal dominant inheritance of the disease trait with complete penetrance and a gene frequency of 0.00001 for the affected allele. The marker allele frequencies used for linkage calculations are given in the Généthon human genetic map public database (ftp://ftp.genethon.fr/pub/Gmap/Nature-1995/alleles/all_chrom6). LOD scores were confirmed independently using the LOD Score Analysis method as defined by Morton.[36]

2.2.5 Haplotype Analysis

Extended haplotypes were constructed manually based on the published sequential order of chromosome 6q microsatellite markers listed in Figure 2.4 and as they appear on the Généthon sex-averaged human genetic linkage map.[37] The population frequency of the disease haplotype was calculated using microsatellite marker allele frequencies found on the previously referenced website.

2.3 RESULTS

2.3.1 Clinical Examination

Of the 27 family members studied (Figure 2.1), 15 were diagnosed as affected based on decreased visual acuity and macular changes observed upon funduscopy.

Fundus examination revealed a variety of progressive abnormalities from retinal pigment epithelial defects (Figure 2.2A) to pattern dystrophy (Figure 2.2C, 2.2D) to a macular atrophic area surrounded by flecks (Figure 2.2E, 2.2F). The dystrophy was further characterized by autosomal dominant inheritance, with affected members of both sexes appearing in each generation (Figure 2.1), and a late age of onset, with symptoms usually appearing in the 4th or 5th decade of life. Both visual acuity and colour discrimination displayed progressive deterioration with increasing age (Figure 2.3). No patient had a significant refractive error. Fluorescein angiography performed for patients III-24 (Figure 2.2B), IV-5, and IV-8 revealed parafoveal window defects and the absence of the diffuse dark choroid typically seen in cases of classical Stargardt macular dystrophy. Both photopic and scotopic ERG waveforms were normal among most of the affected individuals tested (data not shown). Individual V-13 was tested at age 14 according to ISCEV standards and displayed moderately reduced b-wave amplitudes for both the photopic and scotopic response (data not shown), indicative of a cone and rod dysfunction. Dark adaptation curves were normal for all individuals tested except for patient III-15 who displayed an abnormal response (data not shown).

2.3.2 Genetic Analysis

As a genetic locus for autosomal dominant Stargardt-like dystrophy had previously been assigned to human chromosome 6q,[12] microsatellite markers in this region were tested for linkage to the disease locus in the present family. Linkage analysis was performed with 12 polymorphic 6q markers: D6S271, D6S402, D6S286, D6S284, D6S445, D6S1634, D6S1595, D6S1644, D6S300, D6S1716, D6S1717, and D6S475. Two-point linkage analysis yielded a significant positive LOD score of $Zmax$ = 5.50 at a recombination fraction of 0.00 for marker locus D6S300, and a $Zmax$ of 4.24 at a recombination fraction of 0.00 for the adjacent marker D6S1644, thus allowing the formal acceptance of linkage of the disease locus to chromosome 6q (Table 1).

Subsequent haplotype analysis using a total of 43 chromosome 6q microsatellite markers enabled definition of the disease haplotype for the family. The disease haplotype for 12 representative markers that span the region defined by loci D6S271 to D6S475, characteristic of all affected members of the family examined, is shown in Figure 2.4. Based on documented marker allele frequencies and the assumption of Hardy-Weinberg equilibrium, the population frequency of the complete disease haplotype spanning all 43 markers analyzed is calculated as 8.86 x 10^{-24}, suggesting that the probability of this haplotype occurring in the general population by chance is extremely low. The calculated population frequency may represent the minimum probability of randomly observing the combination of marker alleles comprising the disease haplotype due to the possible occurrence of linkage disequilibrium.

Upon examination of the chromosome 6q haplotypes for various members of the pedigree, a number of meiotic recombination events were detected, thus further narrowing the minimum critical region that contains the disease gene. Individuals IV-5, IV-6, IV-8, IV-9, and V-1 are recombinant for marker D6S271, while patients IV-14, IV-17, and V-13 are also recombinant for marker D6S402 (Figure 2.4). This excludes the short arm of chromosome 6 as well as the approximately 7-cM region on the long arm centromeric to D6S402 from harbouring the disease gene. A recombination event was also detected in normal individual V-5, excluding markers telomeric to D6S1716 from the disease haplotype (Figure 2.4). Comparison of the maternally inherited chromosome 6q haplotype of individual V-5 with the corresponding genomic region in her heterozygote affected mother, IV-8, clearly indicates that the alleles for markers D6S1716, D6S1717, and D6S475 were derived from the affected chromosome, while the remainder of the 6q marker alleles analyzed were inherited from the normal maternal chromosome (Figure 2.4). No symptoms of the disease have been observed for individual V-5, who is currently 29 years of age. Taken together, these observations place the disease locus for the family in a region where the loci for other retinal diseases and particularly other autosomal dominant Stargardt-like macular dystrophies have been mapped (Figure 2.5).

2.4 DISCUSSION

This report describes a form of macular degeneration segregating in a large, multi-generation Canadian family. Impairments in visual acuity, central vision, and colour discrimination among affected individuals are consistent with the involvement of the macular region. Funduscopy confirms that atrophy of the macula underlies the abnormalities in visual function. Based on detailed clinical evaluation including fluorescein angiography and electroretinography performed on affected individuals, the disease can be further classified as a Stargardt-like disorder, in which the yellow flecks observed surrounding the atrophic macular region are a prominent feature.[39] However, the combination of clinical features observed in family members diagnosed with the disorder is unique and is collectively unlike any previously described macular dystrophy phenotype. Delineating the disease in the family from classical Stargardt disease are the observations of a relatively late age of onset, the unique appearance of the fundus, and the absence of the dark or obscured choroid typifying fluorescein angiographs of classical Stargardt patients.[39] Furthermore, the autosomal dominant mode of inheritance within the family distinguishes the disorder as genetically distinct from autosomal recessive Stargardt disease. The phenotypic features noted also define this disorder as a novel clinical entity compared with other documented autosomal dominantly inherited Stargardt-like macular dystrophies. Although exhibiting major common features such as atrophic lesions of the macula and retinal pigment epithelium, the presence of fundus flecks, and decreased central visual acuity, notable differences enable categorization of the presently described disease as clinically distinct from STGD2, STGD3 and STGD4 as well as other forms of dominantly inherited macular dystrophy. STGD2 and STGD3 are quite similar in their presentation and differ from the disease documented here most markedly in terms of showing an earlier age of onset (decreased visual acuity is often detected by age 20), variable expressivity, and defects in colour vision observed in some but not all affected individuals.[12-14] Additionally, STGD2 is associated with markedly abnormal ERG responses.[13] STGD4 displays the dark choroid characteristic of autosomal recessive Stargardt disease,[15] which was absent in fluorescein angiography of patients within this family. The disorder described here most closely resembles the adMD phenotype,[27] but again with less variability in clinical presentation. Therefore, the collection of symptoms and clinical features compiled upon evaluation of the patients in the present study further demonstrate the heterogeneity that exists within the group of Stargardt-like disorders.

Similar phenotypic features and the late age of onset of the disorder resembles cases of AMD. This may reflect common pathogenetic mechanisms underlying both diseases. Thus a hereditable macular disease with similar clinical parameters may serve as a model for assessing potential genetic contributions to the development of AMD.

Genetic analysis of the family places the disease locus on the long arm of human chromosome 6. This finding genetically excludes both STGD2 and STGD4 as possible candidate disease loci, as they map to chromosome 13q34 and chromosome 4p respectively.[13,15] Microsatellite marker D6S271, located on human chromosome 6p, was found to be excluded from the common disease haplotype through meiotic recombination events occurring in both affected and normal individuals in the pedigree. This marker is linked to the RDS/peripherin gene, mutations in which have been associated with a number of retinal diseases.[28-33] Therefore, this gene is not a candidate for the disease segregating in the pedigree described here. Markers D6S1716 and D6S1717 are linked proximally to the North Carolina macular dystrophy locus present on chromosome 6q,[22,23] and may putatively be excluded from the disease haplotype based on inheritance of the disease-associated alleles for these markers in a phenotypically normal yet genetically recombinant individual, V-5. V-5 is presently 29 years old and does not have any symptoms of being affected by the disease at this age. This individual will be re-examined on a regular basis to monitor her vision and confirm her unaffected status as she ages into her 4th decade. If V-5 proves to be unaffected, markers D6S1716 and D6S1717 can then be definitively excluded from the disease haplotype, and this would subsequently limit the minimal critical region to a 23-cM interval on the long arm of human chromosome 6 flanked by markers D6S402 and D6S1716. Regardless of this fine tuning, the current localization of the disease locus still significantly overlaps the minimal critical regions documented for other retinal disease loci that map to 6q.

The existence of a number of retinal disease loci in this region of the human genome may suggest a clustering of several different retinal genes on chromosome 6q that, when mutated, give rise to distinct retinal diseases with similar clinical manifestations. This type of locus heterogeneity exists for the cone-rod dystrophies, which exhibit similarities in phenotype, however are associated with mutations in more than 10 different genes.[19,40,41] Alternatively, multiple disorders may result from the occurrence of functionally different mutations within particular retina-expressed genes. The finding that a number of retinal diseases map to the same region on chromosome 6 suggests the possibility that various mutations within a single gene may underlie the myriad of clinical features observed. A recent report documenting an autosomal dominant macular dystrophy sharing features of both cone-rod dystrophy and recessive Stargardt disease and also mapping to chromosome 6q14 suggests possible allelism of the disorder with either CORD7 or STGD3 or both.[16] Indeed, such allelic heterogeneity has been demonstrated for a variety of genetic conditions including retinal dystrophies. For example, different mutations in the RDS/peripherin gene have been shown to give rise to several forms of retinal degeneration including retinitis pigmentosa, cone-rod dystrophy, and pattern dystrophies of the macula.[28-33] Similarly, allelic variants of ABCR gene mutations have been associated with Stargardt disease, retinitis pigmentosa, cone-rod dystrophy, and AMD.[18,41-44]

In summary, the genetic interval defined for the macular disease locus segregating in the Canadian family overlaps that described for other retinal dystrophies including at least two other autosomal dominant Stargardt-like disease loci, STGD3 and adMD. Thus the existence of the disease in this family confirms the map location of these additional loci which may represent the same gene, or alternatively may correspond to a distinct gene in the same genomic region on chromosome 6q.

Efforts to identify the disease gene in this family will prove beneficial for understanding the etiology of this particular disorder and potentially also for other retinal dystrophies mapped to the same chromosomal location. Cloning of the gene and study of its biological function will provide valuable insights into both normal retinal physiology and the pathogenesis of macular degenerative disease. This in turn may lead to the development of novel therapeutic strategies for these severely debilitating eye disorders.

2.5 ACKNOWLEDGEMENTS

The family members and genetic counsellors, Ms. Janelle Schmidt and Ms. Mai Tran, are thanked for their participation in this study. This work was supported by the Alberta Heritage Foundation for Medical Research (PSL, PWW) and the E. A. Baker Foundation of the Canadian National Institute for the Blind (PWW).

2.6 REFERENCES

1. Leibowitz HM, Krueger DE, Maunder LR, Milton RC, Kini MM, Kahn HA, Nickerson RJ, Pool J, Colton TL, Ganley JP, Loewenstein JI, Dawber TR. (1980) The Framingham Eye Study monograph: An ophthalmological and epidemiological study of cataract, glaucoma, diabetic retinopathy, macular degeneration, and visual acuity in a general population of 2631 adults, 1973-1975. *Surv. Ophthalmol.* 24(Suppl):335-610.

2. Egan KM, Seddon JM. (1994) Age-related macular degeneration: Epidemiology. In Albert DM, Jackobiec FA (eds): *Principles and Practice of Ophthalmology: Basic Sciences*. Philadelphia: W B Saunders Company. pp.1266-1274.

3. Mah DY, Wong PW, Edwards A, MacDonald IM. (1998) Recent advances in the genetics of macular dystrophies. *Can. J. Ophthalmol.* 33:135-143.

4. Kaplan J, Bonneau D, Frézal J, Munnich A, Dufier JL. (1990) Clinical and genetic heterogeneity in retinitis pigmentosa. *Hum. Genet.* 85:635-642.

5. Stargardt K. (1909) Über familiäre, progressive Degeneration in der Maculagegend des Auges. *Albrecht von Graefes Arch. Klin. Ophthal.* 71:534-550.

6. Hadden OB, Gass JD. (1976) Fundus flavimaculatus and Stargardt's disease. *Am. J. Ophthalmol.* 82:527-539.

7. Noble KG, Carr RE. (1979) Stargardt's disease and fundus flavimaculatus. *Arch. Ophthalmol.* 97:1281-1285.

8. Eagle RC Jr, Lucier AC, Bernardino VB Jr, Yanoff M. (1980) Retinal pigment epithelial abnormalities in fundus flavimaculatus: a light and electron microscopic study. *Ophthalmology* 87:1189-1200.

9. Lopez PF, Maumenee IH, de la Cruz Z, Green WR. (1990) Autosomal-dominant fundus flavimaculatus: clinicopathologic correlation. *Ophthalmology* 97:798-809.

10. Young RW. (1987) Pathophysiology of age-related macular degeneration. *Surv. Ophthalmol.* 31:291-306.

11. Cibis GW, Morey M, Harris DJ. (1980) Dominantly inherited macular dystrophy with flecks (Stargardt). *Arch. Ophthalmol.* 98:1785-1789.

12. Stone EM, Nichols BE, Kimura AE, Weingeist TA, Drack A, Sheffield VC. (1994) Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. *Arch. Ophthalmol.* 112:765-772.

13. Zhang K, Bither PP, Park R, Donoso LA, Seidman JG, Seidman CE. (1994) A dominant Stargardt's macular dystrophy locus maps to chromosome 13q34. *Arch. Ophthalmol.* 112:759-764.

14. Edwards AO, Miedziak A, Vrabec T, Verhoeven J, Acott TS, Weleber RG, Donoso LA. (1999) Autosomal dominant Stargardt-like macular dystrophy: I. Clinical characterization, longitudinal follow-up, and evidence for a common ancestry in families linked to chromosome 6q14. *Am. J. Ophthalmol.* 127:426-435.

15. Kniazeva M, Chiang MF, Morgan B, Anduze AL, Zack DJ, Han M, Zhang K. (1999) A new locus for autosomal dominant stargardt-like disease maps to chromosome 4. *Am. J. Hum. Genet.* 64:1394-1399.

16. Kniazeva MF, Chiang MF, Cutting GR, Zack DJ, Han M, Zhang K. (1999) Clinical and genetic studies of an autosomal dominant cone-rod dystrophy with features of Stargardt disease. *Ophthalmic Genet.* 20:71-81.

17. Kaplan J, Gerber S, Larget-Piet D, Rozet JM, Dollfus H, Dufier JL, Odent S, Postel-Vinay A, Janin N, Briard ML, Frézal J, Munnich A. (1993) A gene for Stargardt's disease (fundus flavimaculatus) maps to the short arm of chromosome 1. *Nature Genet.* 5:308-311.

18. Allikmets R, Singh N, Sun H, Shroyer NF, Hutchinson A, Chidambaram A, Gerrard B, Baird L, Stauffer D, Peiffer A, Rattner A, Smallwood P, Li Y, Anderson KL, Lewis RA, Nathans J, Leppert M, Dean M, Lupski JR. (1997) A photoreceptor cell-specific ATP-binding transporter gene (*ABCR*) is mutated in recessive Stargardt macular dystrophy. *Nature Genet.* 15:236-246.

19. Kelsell RE, Gregory-Evans K, Gregory-Evans CY, Holder GE, Jay MR, Weber BH, Moore AT, Bird AC, Hunt DM. (1998) Localization of a gene (CORD7) for a dominant cone-rod dystrophy to chromosome 6q. *Am. J. Hum. Genet.* 63:274-279.

20. Ruiz A, Borrego S, Marcos I, Antinolo G. (1998) A major locus for autosomal recessive retinitis pigmentosa on 6q, determined by homozygosity mapping of chromosomal regions that contain gamma-aminobutyric acid-receptor clusters. *Am. J. Hum. Genet.* 62:1452-1459.

21. Khaliq S, Hameed A, Ismail M, Mehdi SQ, Bessant DA, Payne AM, Bhattacharya SS. (1999) Refinement of the locus for autosomal recessive Retinitis pigmentosa (RP25) linked to chromosome 6q in a family of Pakistani origin. *Am. J. Hum. Genet.* 65:571-574.

22. Small KW, Weber JL, Roses A, Lennon F, Vance JM, Pericak-Vance MA. (1992) North Carolina macular dystrophy is assigned to chromosome 6. *Genomics* 13:681-685.

23. Sauer CG, Schworm HD, Ulbig M, Blankenagel A, Rohrschneider K, Pauleikhoff D, Grimm T, Weber BH. (1997) An ancestral core haplotype defines the critical region harbouring the North Carolina macular dystrophy gene (MCDR1). *J. Med. Genet.* 34:961-966.

24. Kelsell RE, Godley BF, Evans K, Tiffin PA, Gregory CY, Plant C, Moore AT, Bird AC, Hunt DM. (1995) Localization of the gene for progressive bifocal chorioretinal atrophy (PBCRA) to chromosome 6q. *Hum. Mol. Genet.* 4:1653-1656.

25. Dharmaraj S, Li Y, Robitaille JM, Silva E, Zhu D, Mitchell TN, Maltby LP, Baffoe-Bonnie AB, Maumenee IH. (2000) A novel locus for Leber congenital amaurosis maps to chromosome 6q. *Am. J. Hum. Genet.* 66:319-326.

26. Kniazeva M, Traboulsi EI, Yu Z, Stefko ST, Gorin MB, Shugart YY, O'Connell JR, Blaschak CJ, Cutting G, Han M, Zhang K. (2000) A new locus for dominant drusen and macular degeneration maps to chromosome 6q14. *Am. J. Ophthalmol.* 130:197-202.

27. Griesinger IB, Sieving PA, Ayyagari R. (2000) Autosomal dominant macular atrophy at 6q14 excludes CORD7 and MCDR1/PBCRA loci. *Invest. Ophthalmol. Vis. Sci.* 41:248-255.

28. Nichols BE, Sheffield VC, Vandenburgh K, Drack AV, Kimura AE, Stone EM. (1993) Butterfly-shaped pigment dystrophy of the fovea caused by a point mutation in codon 167 of the *RDS* gene. *Nature Genet.* 3:202-207.

29. Wells J, Wroblewski J, Keen J, Inglehearn C, Jubb C, Eckstein A, Jay M, Arden G, Bhattacharya S, Fitzke F, Bird A. (1993) Mutations in the human retinal degeneration slow (*RDS*) gene can cause either retinitis pigmentosa or macular dystrophy. *Nature Genet.* 3:213-218.

30. Weleber RG, Carr RE, Murphey WH, Sheffield VC, Stone EM. (1993) Phenotypic variation including retinitis pigmentosa, pattern dystrophy, and fundus flavimaculatus in a single family with a deletion of codon 153 or 154 of the *peripherin/RDS* gene. *Arch. Ophthalmol.* 111:1531-1542.

31. Nakazawa M, Naoi N, Wada Y, Nakazaki S, Maruiwa F, Sawada A, Tamai M. (1996) Autosomal dominant cone-rod dystrophy associated with a Val200Glu mutation of the *peripherin/RDS* gene. *Retina* 16:405-410.

32. Nakazawa M, Kikawa E, Chida Y, Wada Y, Shiono T, Tamai M. (1996) Autosomal dominant cone-rod dystrophy associated with mutations in codon 244 (Asn244His) and codon 184 (Tyr184Ser) of the *peripherin/RDS* gene. *Arch. Ophthalmol.* 114:72-78.

33. Kohl S, Christ-Adler M, Apfelstedt-Sylla E, Kellner U, Eckstein A, Zrenner E, Wissinger B. (1997) *RDS/peripherin* gene mutations are frequent causes of central retinal dystrophies. *J. Med. Genet.* 34:620-626.

34. International Standardization Committee. (1989) Standard for clinical electroretinography. *Arch. Ophthalmol.* 107:816-819.

35. Lathrop GM, Lalouel JM. (1984) Easy calculations of lod scores and genetic risks on small computers. *Am. J. Hum. Genet.* 36:460-465.

36. Morton NE. (1955) Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* 7:277-318.

37. Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J. (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380:152-154.

38. Westheimer G. (1992) Visual acuity. In Hart WM Jr (ed): *Adler's Physiology of the Eye: Clinical Application*. St. Louis: Mosby-Year Book. pp.531-547.

39. Gass JDM. (1987) Stargardt's disease (fundus flavimaculatus). In Gass JDM (ed): *Stereoscopic Atlas of Macular Diseases: Diagnosis and Treatment*. 3rd ed. St. Louis: C V Mosby. pp.256-261.

40. Simunovic MP, Moore AT. (1998) The cone dystrophies. *Eye* 12:553-565.

41. Cremers FP, van de Pol DJ, van Driel M, den Hollander AI, van Haren FJ, Knoers NV, Tijmes N, Bergen AA, Rohrschneider K, Blankenagel A, Pinckers AJ, Deutman AF, Hoyng CB. (1998) Autosomal recessive retinitis pigmentosa and cone-rod dystrophy caused by splice site mutations in the Stargardt's disease gene *ABCR*. *Hum. Mol. Genet.* 7:355-362.

42. Martinez-Mir A, Paloma E, Allikmets R, Ayuso C, del Rio T, Dean M, Vilageliu L, Gonzalez-Duarte R, Balcells S. (1998) Retinitis pigmentosa caused by a homozygous mutation in the Stargardt disease gene *ABCR*. *Nature Genet.* 18:11-12.

43. Allikmets R, Shroyer NF, Singh N, Seddon JM, Lewis RA, Bernstein PS, Peiffer A, Zabriskie NA, Li Y, Hutchinson A, Dean M, Lupski JR, Leppert M. (1997) Mutation of the Stargardt disease gene (*ABCR*) in age-related macular degeneration. *Science* 277:1805-1807.

44. Van Driel MA, Maugeri A, Klevering BJ, Hoyng CB, Cremers FP. (1998) ABCR unites what ophthalmologists divide(s). *Ophthalmic Genet.* 19:117-122.

Figure 2.1 Pedigree for the Canadian family showing autosomal dominant inheritance of a Stargardt-like macular dystrophy. Individuals are identified by generation and pedigree number. Solid symbols indicate affected family members while unaffected or unexamined individuals are represented by open symbols. Circles represent females, squares represent males, and diamonds represent individuals of unknown gender. Numbers within the symbols indicate the number of individuals known to exist of the specified gender (N = unknown number). Diagonal lines through symbols indicate deceased individuals. Pedigree assembled by Dr. Ian M. MacDonald and Dr. William G. Pearce (University of Alberta).

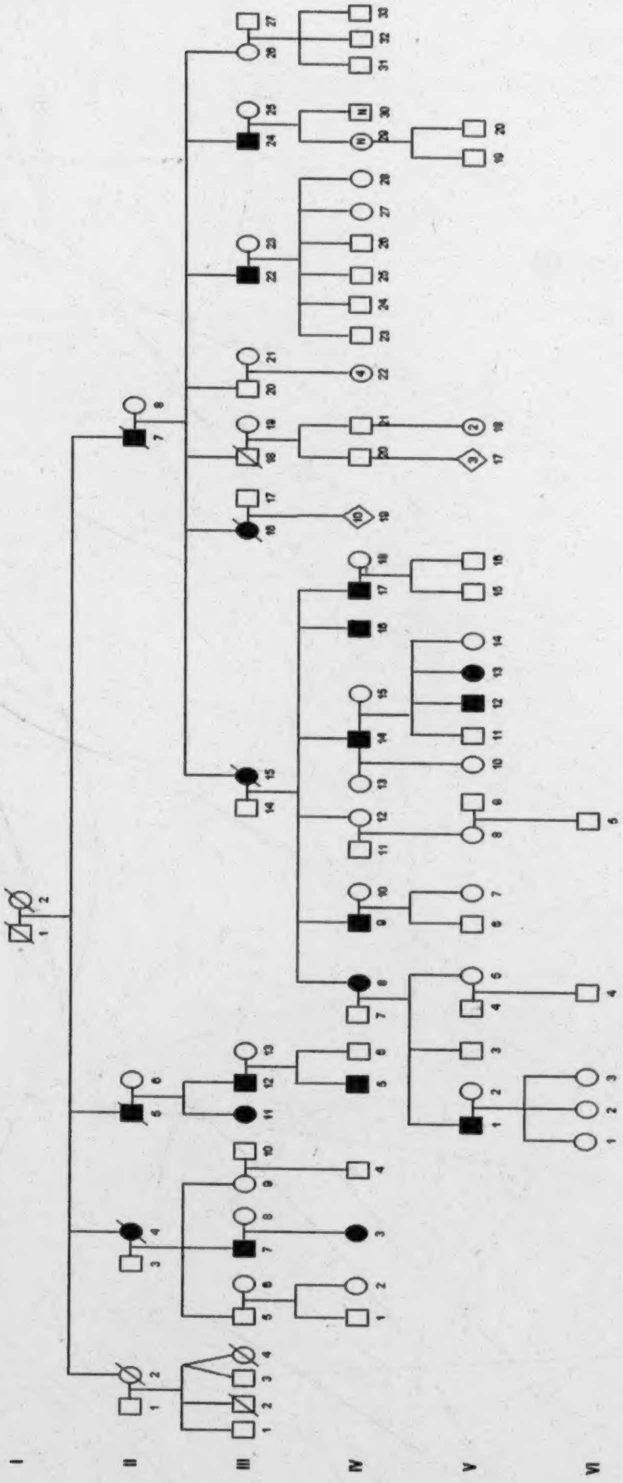

I
II
III
IV
V
VI

Figure 2.2 Funduscopy and fluorescein angiography for two affected individuals within the Canadian autosomal dominant Stargardt-like macular dystrophy pedigree. Fundus photograph (Panel A) and fluorescein angiogram (Panel B) of the right eye of patient III-24 at age 39. Visual acuity was 20/30. Patient III-24 was re-examined at age 48 (Panel C) with visual acuity at 20/50 (OD), and again at age 59 (Panel D) with visual acuity at 20/80 (OD). Fundus photographs are also shown for patient IV-14 at age 21 (Panel E) with visual acuity at 20/200 (OD), and for re-examination at age 43 (Panel F) with visual acuity at 20/400 (OD). Images provided by Dr. Ian M. MacDonald.

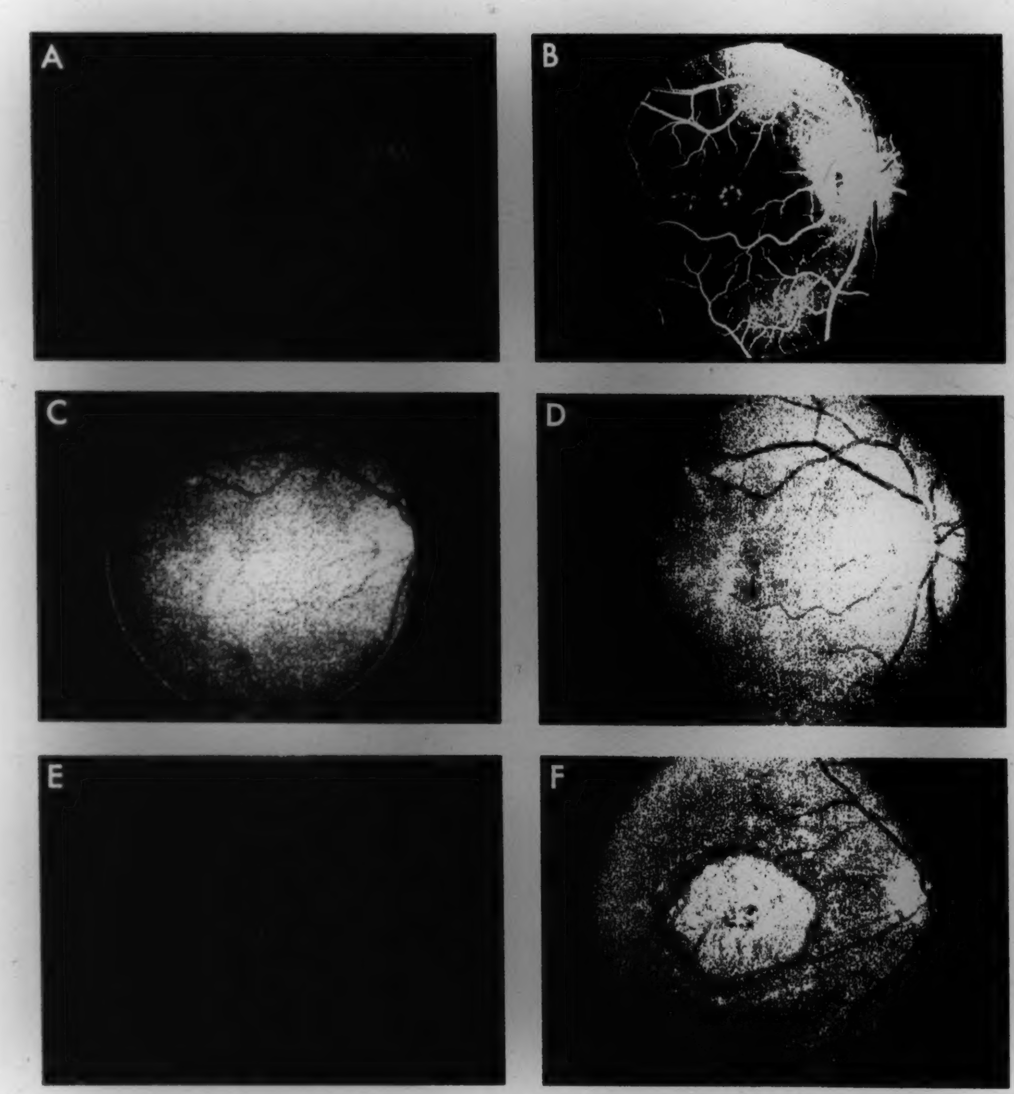
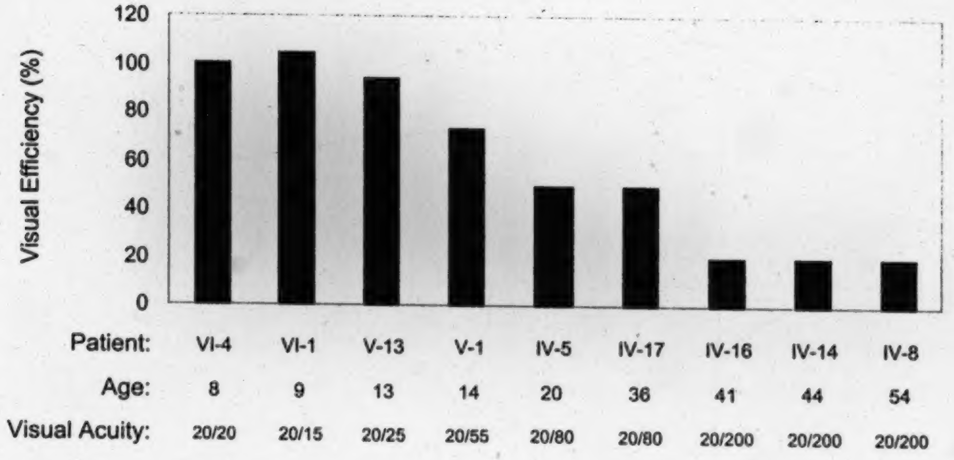

A
B
C
D
E
F

Figure 2.3 Summary of visual acuity (Panel A) and colour vision (Panel B) tests for different individuals within the Canadian autosomal dominant Stargardt-like macular dystrophy family. Snell-Sterling's visual efficiency ratings[17] are determined based on visual acuity measurements shown below the graph for each individual. Progressive deterioration is observed with increasing age of the patients. Ophthalmologic examinations performed by Dr. Ian M. MacDonald.

A


Visual Efficiency (%)
120
100
80
60
40
20
0
Patient:
VI-4
VI-1
V-13
V-1
IV-5
IV-17
IV-16
IV-14
IV-8
Age:
8
9
13
14
20
36
41
44
54
Visual Acuity:
20/20
20/15
20/25
20/55
20/80
20/80
20/200
20/200
20/200

B

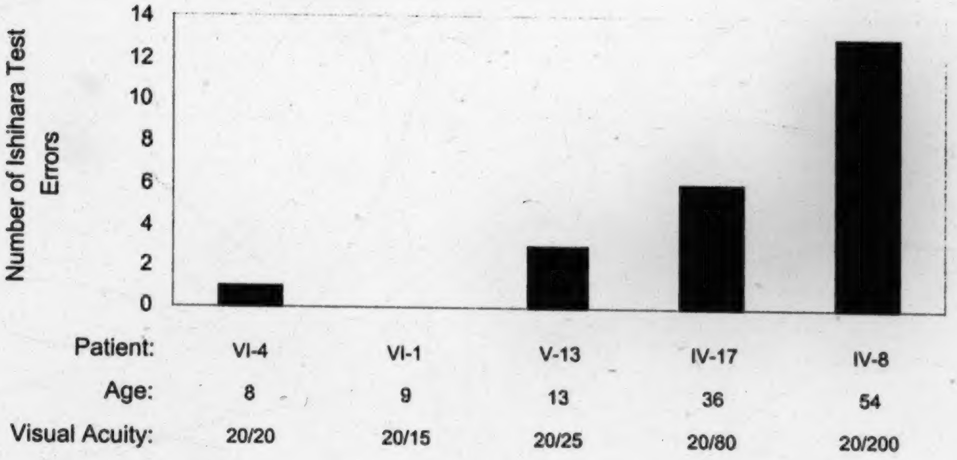

Number of Ishihara Test Errors
14
12
10
8
6
4
2
0
Patient:
VI-4
VI-1
V-13
IV-17
IV-8
Age:
8
9
13
36
54
Visual Acuity:
20/20
20/15
20/25
20/80
20/200

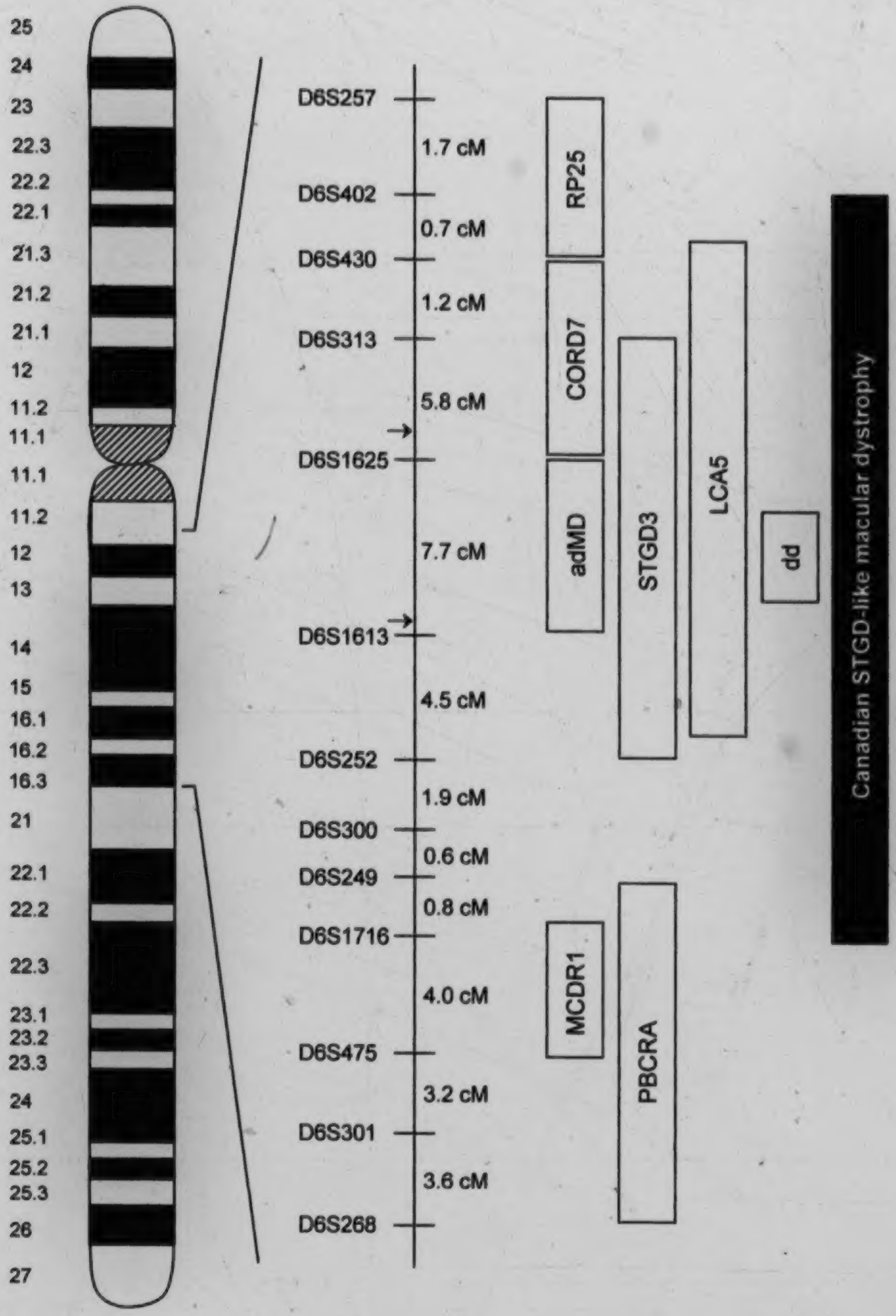
Figure 2.4 Chromosome 6q microsatellite marker haplotypes for individuals within the Canadian autosomal dominant Stargardt-like macular dystrophy pedigree. The numbers within the boxed areas represent the alleles detected in family members for the corresponding polymorphic microsatellite markers listed on the left. Dashes (-) denote that data was not available. Dark portions of the haplotypes represent the alleles segregating with the disease phenotype. Dark symbols above the boxed haplotypes indicate affected individuals, while open symbols indicate normal individuals; the pedigree identification for each individual is given within the symbols. [a]affected chromosome 6q haplotype; [b]normal chromosome 6q haplotype.

Polymorphic Marker	III-7	IV-3	III-11	III-12	IV-5	IV-6	IV-8^a	IV-8^b	V-1	V-5	IV-9	IV-14	V-13	IV-17
D6S271	4	4	4	4	5	4	5	1	5	1	5	5	5	5
D6S402	5	5	-	5	5	9	5	5	5	5	5	9	9	9
D6S286	4	4	4	4	4	4	4	12	4	12	4	4	4	4
D6S284	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D6S445	3	3	3	3	3	3	3	3	3	3	3	3	3	3
D6S1634	4	4	4	4	4	3	4	3	4	3	4	4	4	4
D6S1595	2	2	2	2	-	3	2	3	2	3	2	2	2	2
D6S1644	3	3	3	3	3	2	3	5	3	5	3	3	3	3
D6S300	10	10	10	10	10	5	10	2	10	2	10	10	10	10
D6S1716	2	2	2	2	-	2	2	1	2	2	2	2	-	2
D6S1717	1	1	1	1	1	4	1	4	1	1	1	1	1	1
D6S475	7	7	7	7	7	7	7	6	7	7	7	7	7	7

Table 2.1 Two-point LOD scores for linkage between the autosomal dominant Stargardt-like macular dystrophy locus within the Canadian pedigree and microsatellite markers on chromosome 6q.

Marker	Two-point LOD scores at theta =						*Zmax*	*θmax*
	0.00	0.05	0.10	0.20	0.30	0.40		
D6S271	-∞	0.49	0.86	0.95	0.73	0.37	0.96	0.18
D6S402	-∞	0.19	0.49	0.50	0.30	0.11	0.54	0.15
D6S286	0.86	0.72	0.59	0.36	0.18	0.06	0.86	0.00
D6S284	1.02	0.87	0.73	0.46	0.25	0.10	1.02	0.00
D6S445	1.88	1.66	1.42	0.95	0.50	0.17	1.88	0.00
D6S1634	2.57	2.26	1.95	1.33	0.75	0.26	2.57	0.00
D6S1595	1.55	1.35	1.16	0.80	0.49	0.23	1.55	0.00
D6S1644	4.24	3.84	3.42	2.52	1.58	0.69	4.24	0.00
D6S300	5.50	5.03	4.54	3.47	2.31	1.11	5.50	0.00
D6S1716	-∞	0.65	0.79	0.73	0.51	0.23	0.80	0.13
D6S1717	-∞	1.77	1.81	1.55	1.10	0.55	1.81	0.08
D6S475	-∞	0.65	0.79	0.73	0.51	0.23	0.80	0.13

Figure 2.5 Chromosome 6 ideogram showing the relative positions of the retinal disease loci that have been mapped to 6q. Genetic distances between microsatellite marker positions (not to scale) correspond to sex-averaged distances given in the Généthon human genetic map public database (see Methods). Filled rectangle represents the existing minimal critical region for the disease locus identified in this study. Arrows indicate positions of microsatellite markers D6S1596 and D6S1644 that define the boundaries of the chromosome 6q interval analyzed for candidate retinal disease genes (see Chapter 3 for details). *dd,* dominant drusen with macular degeneration.[26]


25
24
23
22.3
22.2
22.1
21.3
21.2
21.1
12
11.2
11.1
11.1
11.2
12
13
14
15
16.1
16.2
16.3
21
22.1
22.2
22.3
23.1
23.2
23.3
24
25.1
25.2
25.3
26
27
D6S257
1.7 cM
D6S402
0.7 cM
D6S430
1.2 cM
D6S313
5.8 cM
D6S1625
7.7 cM
D6S1613
4.5 cM
D6S252
1.9 cM
D6S300
0.6 cM
D6S249
0.8 cM
D6S1716
4.0 cM
D6S475
3.2 cM
D6S301
3.6 cM
D6S268
RP25
CORD7
adMD
MCDR1
STGD3
PBCRA
LCA5
pd
Canadian STGD-like macular dystrophy

Chapter 3

A Tissue-Selective Transcript Map for a Retinal Disease Locus-Rich Region of Human Chromosome 6q

3.1 INTRODUCTION

The long arm of human chromosome 6 is a genomic region that is particularly rich in retinal disease loci. At least 8 genetic loci for inherited retinal dystrophies have been mapped to this region. These include autosomal dominant macular dystrophy (STGD3,[1-3] adMD[4]), autosomal recessive retinitis pigmentosa (RP25),[5,6] cone-rod dystrophy (CORD7),[7] North Carolina macular dystrophy (MCDR1),[8] progressive bifocal chorioretinal atrophy (PBCRA),[9] Leber congenital amaurosis (LCA5),[10] dominant drusen,[11] and benign concentric annular macular dystrophy (BCAMD).[12] All of these disorders involve some type of progressive photoreceptor and/or retinal pigment epithelial cell degeneration that ultimately leads to permanent and irreversible blindness in affected individuals. Furthermore, chromosomal aberrations such as deletions, duplications, and translocations involving telomeric regions of chromosome 6q have been associated with additional retinal abnormalities.[13-18] The genes responsible for the majority of these chromosome 6q-linked retinal disorders remain as yet unknown.

Traditional strategies for identifying candidate disease genes towards elucidating the genetic basis of Mendelian disorders in the pre-genome era typically involved generating physical maps of genetic intervals harbouring disease loci, which required the acquisition of genomic clones spanning an interval of interest, laborious assembly procedures to deduce and confirm overlaps between adjacent clones, and sequencing to verify the junctions between the constituent pieces of DNA. Following creation of the physical map, the linear DNA sequence assembled was analyzed for the presence of putative exons and protein-encoding regions to locate and define potential genes that were considered as candidate disease targets. The process was further hindered by the existence of sequence gaps in the contig generated due to inaccessibility of genomic interval-spanning clones, as well as the lack of a means of prioritizing regions in which to look for sequences of particular interest.

The recent completion of the Human Genome Sequencing Project has greatly facilitated this cumbersome endeavour by providing not only the entire genetic sequence of each human chromosome, but also yielding information pertaining to gene expression through the generation and analysis of expressed sequence tags (ESTs). The subsequent elucidation of genome sequences for other species has further enabled an assessment of gene conservation through evolution by comparative genome sequence analysis. As a result of the availability of new genomic tools and resources, both the positional cloning (reverse genetics) and candidate gene (forward genetics) approaches for identifying novel disease genes have been made more efficient. The identification of previously unknown marker loci has enabled the generation of high density genetic maps that can be used for genomic mapping and linkage studies, thereby facilitating the cloning of disease genes based on their chromosomal position. The isolation of ESTs from various tissues and predicted protein-encoding sequences in genomic DNA from which novel genes can be defined provides a new pool of molecular candidates that can be screened for disease-causing mutations in affected individuals. Information on tissue expression profiles and evolutionary conservation of such genes can be used to limit the search parameters to isolate candidates of particular interest that may represent likely targets of disease.

Making use of all available genetic resources, a combined "positional-candidate" strategy is now commonly used to identify disease genes. Such an approach was adopted in the successful search for the genes underlying two forms of autosomal dominant macular degeneration and cone-rod dystrophy mapping to chromosome 6q. The STGD3/adMD gene, elongation of very long-chain fatty acids 4 (*ELOVL4*), was isolated as a result of positional cloning, EST analysis, and exon prediction within the genomic sequence comprising the minimal critical interval found to harbour the disease locus (see Chapter 5 for details).[19,20] The CORD7 gene, designated Rab3A-interacting molecule (*RIM1*), was identified through linkage mapping followed by mutation screening of candidate genes that were defined by their localization to the critical interval and a demonstrated retinal function or expression.[21] To date, these are the only known retinal genes that have been characterized in this region of the genome in which mutations associated specifically with retinal disease have been found. Very recently, linkage-based studies identified mutations in the peroxisome biogenesis gene peroxin 7 (*PEX7*) on chromosome 6q21-q22.2 in patients with Refsum disease, a multi-symptom disorder characterized in part by progressive adult retinitis pigmentosa.[22] Another peroxin, *PEX3*, maps to 6q23-q24 and has been identified as a gene underlying some cases of Zellweger Syndrome,[23,24] a condition in which retinal defects are consistently observed.[25]

Also mapping to chromosome 6q are other genes that have been shown to be expressed in and/or play a functional role with respect to retinal cell types, including interphotoreceptor matrix proteoglycan 1 (*IMPG-1*),[26] gamma-aminobutyric acid receptor subunits *GABRR1* and *GABRR2*,[27,28] phosphatase of regenerating liver 1 (*PRL-1*),[29] and chromosome 6 open reading frame 37 (*C6orf37*).[30] However, no definitive disease-causing mutations associated with any retinal disorder have been described for these genes, and some of them have in fact been excluded as candidates in a number of clinical and genetic studies.[30-33]

The aim of the present study is to uncover additional candidate genes for the retinal diseases that map to chromosome 6q, a veritable 'hotspot' for retinal dysfunction. The method described for classification of candidate disease genes involves combining positional information with gene expression and conservation data. The ability of a given tissue to perform its biological function is dependent on the collection of genes that are normally expressed within that tissue. As a result, any tissue in which a particular gene is expressed may become susceptible to disease in the event of gene dysfunction. Thus the characterization of the expression status of genes in a tissue type of interest identifies possible candidate genes for diseases affecting that tissue. As both the retina and RPE are tissues relevant to the etiology of retinal dystrophy, the expression of putative genes in both of these sub-ocular regions are studied. Sequence conservation of genes and their encoded proteins across multiple and diverse species is indicative of a critical biological function for these molecules that has been maintained over time despite evolutionary pressures. The more species in which homologues of a particular gene are found, the older is the shared sequence, and the more fundamental its function is believed to be. Therefore by assessing the extent of conservation throughout evolution, genes that are more likely to lead to biological dysfunction when mutated can be identified.

Given the density of expressed sequence tags (ESTs) that have been mapped to chromosome 6q, the increasing number of both defined and predicted genes localized to this region, and the existence of numerous retinal disease loci that map to 6q, it is probable that a significant number of as yet uncharacterized genes reside in this area of the genome and that some of these are critical for normal retinal function. In the interest of defining such genes, a retinal and RPE transcript map has been generated for an approximately 8-centimorgan interval of human chromosome 6q flanked by microsatellite markers D6S1596 and D6S1644, an interval which contains or overlaps the disease loci for a number of retinal dystrophies described. By creating a comprehensive, ordered list of all known genetic marker loci found in current human genomic sequence databases and incorporating information obtained from expression analysis of resident ESTs, an integrated physical and transcript map has essentially been developed that can be used as a valuable resource in the search for additional candidate genes on chromosome 6q12-q16 which may underlie the retinal disorders that map to this genomic region.

3.2 MATERIALS AND METHODS

3.2.1 Defining the Chromosome 6q12-q16 Genomic Interval of Interest

Previous linkage analysis of a STGD3 pedigree delineated a 23-cM genomic interval flanked by microsatellite markers D6S402 and D6S1716 on chromosome 6q to contain the disease locus, a region to which other retinal disease loci had also been mapped (Figure 2.5).[3] Further refinement of the critical interval based on the assessment of additional microsatellite markers within recombinant individuals of the pedigree in addition to members of a related adMD family[4] narrowed the genomic region of interest to roughly 8 cM between markers D6S1596 and D6S1644 located within cytogenetic bands 6q12 and 6q16 respectively (Figure 2.5).

3.2.2 Identification and Expression Analysis of ESTs

As described in Chapter 4, section 4.2.1, ESTs localized to the genomic interval of interest on chromosome 6q were identified through searches of public human genome sequence databases[20,30] and were commercially obtained (Research Genetics, Inc.). *In silico* expression profiles of the 139 non-overlapping ESTs selected were provided in the UniGene database (http://www.ncbi.nlm.nih.gov/UniGene/) and accessed using the EST GenBank accession number for each EST as a query term. Sub-retinal and RPE expression of each EST was experimentally assessed through dot blot analysis using macroarrays of the ESTs and total cDNA probes derived from human peripheral retina, fovea, and RPE tissues as described in Chapter 4, section 4.2.1.[20,30] Relative hybridization levels of the ESTs with the different total cDNA probes were assigned by visual comparison of autoradiographic intensities resulting from exposure of the probed macroarrays to X-ray film.

3.2.3 *Construction of the Retinal and RPE Transcript Map*

Nucleotide positions of microsatellite markers D6S1596 and D6S1644 on chromosome 6 as found in the UniSTS database (http://www.ncbi.nlm.nih.gov/UniSTS/) were used as the boundaries of the region of interest in which to locate additional markers, known genes, predicted genes, and pseudogenes included in Build 34 version 1 of the National Center for Biotechnology Information (NCBI)'s assembly of human genomic sequence data represented in the human Map Viewer database (http://www.ncbi.nlm.nih.gov/mapview/). Chromosome 6 nucleotide positions of the ESTs obtained and analyzed for retinal and RPE expression were either found directly in the UniSTS database or were deduced through comparison of the relative position of ESTs in question along the human chromosome 6 genomic contig clone (GenBank accession number NT_007299) that spans the interval of interest as identified through a BLAST search of the existing human genome sequence (http://www.ncbi.nlm.nih.gov/genome/seq/page.cgi?F=HsBlast.html&&ORG=Hs). The ESTs were then accordingly placed in sequence among the list of genes and markers assembled. The map boundaries were extended both centromerically and telomerically to accommodate the ESTs that were originally identified within the D6S1596 to D6S1644 interval in the GeneMap'99 human genome database (http://www.ncbi.nlm.nih.gov/genemap99) but were subsequently relocalized to lie outside of the interval according to the updated human Map Viewer genome resource. The analyzed ESTs shown to currently lie in these extended regions of the interval of interest, i.e. between markers D6S455 - D6S1596 and D6S1644 - D6S1704, do not represent complete retinal or RPE transcript coverage in these areas of chromosome 6q. Nucleotide assignments of ESTs belonging to the same cluster and/or representing the same gene as classified in the UniGene database were integrated and included as a single entry in the map. Average hybridization level scores were calculated and reported for single map entries corresponding to multiple ESTs.

3.2.4 *Expression Profiling and Evolutionary Conservation of Putative Retinal/RPE Genes*

Individual ESTs of interest were further investigated for their expression profiles over a broad range of human tissue types by northern blot analysis and for sequence conservation across different species through Southern zooblot analysis as previously described.[30] Possible homologous and/or orthologous genes predicted by comparison of transcript or annotated genome sequences between species were identified using the NCBI HomoloGene resource (http://www.ncbi.nlm.nih.gov/HomoloGene/). Homologues are defined by this program as gene sequences sharing greater than 70% identity at the nucleotide level with a minimum alignment of 100 base pairs.

3.3 RESULTS

3.3.1 Identification of Retina- or RPE-Expressed 6q12-q16 ESTs

A genomic interval of interest was established through prior genetic studies of macular dystrophy-affected pedigrees[3,4] and corresponded to an approximately 8-cM region of human chromosome 6q. An interval-wide screen of all the human ESTs known to lie between the flanking microsatellite markers D6S1596 and D6S1644 was initiated. A total of 196 ESTs representing 139 non-overlapping transcribed sequences were spotted onto multiple nylon membranes to create identical macroarrays that were subsequently probed with radiolabeled total cDNA pools derived from one of three human ocular tissues: peripheral retina, fovea, and RPE. As these tissues are directly implicated in retinal degenerative disorders and the sites of preferential expression of over two-thirds of all retinal degeneration-associated genes identified to date,[34] ESTs that repeatedly tested positive for hybridization with any of these three probes were considered to represent putative candidate genes for retinal disease.

A sample of the autoradiographic signals generated from the macroarrays is shown in Figure 3.1. The majority of the ESTs analyzed exhibited some degree of retinal and/or RPE expression, with 94% displaying positive hybridization to at least one of the three total cDNA probes, corresponding to 131/139 of the non-overlapping ESTs originally assessed. Most ESTs were found to be expressed at roughly equivalent levels amongst all three tissue types (e.g. Figure 3.1, clones a, d, l). However, differential expression of some of the ESTs was also observed, including those categorized as retina-enriched (e.g. Figure 3.1, clones g, k), fovea-enriched (e.g. Figure 3.1, clone h), and RPE-enriched (e.g. Figure 3.1, clone o) relative to their expression status in the other ocular tissues tested. Validity of the macroarray results was established by observing expected hybridization signals for cDNA clones corresponding to tissue-specific or tissue-predominant transcripts, i.e. rhodopsin cDNA hybridization to the peripheral retina cDNA probe and RPE65 cDNA hybridization to the RPE cDNA probe (Figure 3.1).

The autoradiographic signals arising from the probed macroarrays were assigned numerical values representing relative levels of hybridization intensity in order to create a priority list for further investigation of the individual ESTs. Examples of typical hybridization scores assigned can be seen in the RPE total cDNA-probed panel of dot blots shown in Figure 3.1: hybridization level 1 – clone e, level 2 – clone h, level 3 – clone g, level 4 – clone o, level 5 – clone a.

3.3.2 Categorization of Ocular Genes Represented by 6q12-q16 ESTs

The sequence-confirmed 131 retina- and/or RPE-expressed 6q12-q16 ESTs were found to correspond to 72 independent genes or EST clusters according to the UniGene database. Examination of the *in silico* expression profiles for these known or putative genes, also given in the UniGene database, revealed that the majority of the genes had previously not been known to be expressed in the sub-retinal or RPE tissues tested (Figure 3.2). Of the genes represented by ESTs exhibiting retinal (either peripheral retina or fovea) expression on the macroarrays, 38% (27/72) had constituent ESTs derived from a catalogued retinal cDNA library, while 62% (45/72) correspond to "novel" retina-expressed genes. Similarly, 31% (22/72) of the fovea-expressed genes had previously been represented in human foveal libraries reported in UniGene, while 69% (50/72) were newly found to be foveal genes in the present study. Previously known RPE genes constituted 33% (24/72) of the EST clusters found to be RPE-expressed upon screening of the macroarrays, while 67% (48/72) were novel RPE-expressed genes. These results indicate that the methodology utilized does indeed effectively lead to the identification of a significant number of new retinal and RPE-expressed genes.

Comparison of the average hybridization scores and thus expression levels for the ESTs comprising the known or putative genes among the three ocular tissues examined enabled categorization of some of the genes as tissue-enriched (Figure 3.2). Based on a higher score assigned to retinal (either peripheral retina or fovea) probe-generated hybridization signals versus RPE probe-generated signals, roughly 36%

(26/72) of the genes represented by the ESTs screened were classified as retina-enriched. Examination of sub-retinal expression status revealed that approximately 8% (6/72) of the genes were found to be enriched in the peripheral retina and 14% (10/72) displayed fovea-enriched expression. RPE-enriched genes represented roughly 14% (10/72) of the 6q12-q16 EST clusters analyzed. The UniGene entries for some of the genes listed other ocular tissues or whole eye sources for constituent ESTs. If total ocular expression is considered in this type of classification of the genes corresponding to the retinal or RPE-expressed ESTs, approximately 54% (39/72) were previously known to be ocular genes while 46% (33/72) were not represented in libraries derived from any eye tissue. Thus in addition to the identification of a number of new ocular genes, a greater proportion of genes expressed in specific sub-ocular tissues than that previously known was revealed through this analysis.

3.3.3 A Comprehensive Retinal and RPE Transcript Map on 6q12-q16

All of the retina- and/or RPE-expressed ESTs analyzed were listed in succession and arranged among the predetermined order of established genes, predicted genes, pseudogenes, and microsatellite markers found in the 6q12-q16 genomic interval of interest based on their reported nucleotide positions on chromosome 6. The list generated essentially comprises a retina and RPE transcript map of this genomic region (Table 3.1). Some of the ESTs that were originally described to lie between markers D6S1596 and D6S1644 were later found to be localized outside this 8-cM (equivalent to 15.3 megabases) interval upon generation of more detailed and accurate versions of the human genome map. As a result, these ESTs currently lie in genomic regions that flank the core critical interval initially defined. The number of non-overlapping retinal or RPE ESTs correctly localized to the original D6S1596 - D6S1644 interval is 110, or 84% of the total number (131) placed on the map. Therefore the map depicted in Table 3.1 contains ESTs that span a 61.6-cM or 72.3-megabase region on human chromosome 6q, thereby extending the interval in which new ocular genes have been defined and providing potential molecular candidates for retinal diseases that map to a broader chromosomal segment.

3.3.4 Further Characterization of Retina/RPE-Expressed 6q12-q16 Genes

Individual ESTs that were selected for further characterization based on their hybridization scores across the retinal and RPE tissues tested were subjected to northern blot analysis to confirm the ocular expression observed on the macroarrays and also to assess their expression profiles in a wide variety of both adult and fetal non-ocular tissue types. The results of such analysis of six of the ESTs examined are shown in Figure 3.3. When probed with the individual EST clones, the northern blots containing human retina- and RPE-derived RNA demonstrated relative expression levels that were consistent with the original macroarray screen results. Ocular-specific expression was not observed for any of the ESTs analyzed, as hybridization signals also appeared for multiple non-ocular RNA samples in all cases. However, differential expression was noted for each of the corresponding genes among the tissue types analyzed, and in both adult and fetal samples, indicating that the ESTs examined do not represent housekeeping genes but instead likely play tissue-selective roles. As expected, for ESTs representing now known genes, both transcript sizes and expression profiles in tissues previously examined were generally consistent with prior reports.[30,35,36]

To investigate the evolutionary conservation of the genes represented by particular ESTs, Southern blots of genomic DNA isolated from a panel of evolutionary divergent species were probed (Figure 3.4). The genes represented by EST clones 1a11 and 1e1 appear to be somewhat restricted in their conservation across species, generating discrete signals for only one or two organisms aside from those seen for the human DNA samples. In contrast, EST clone 1c6 displayed sequence conservation among a more broad range of species, with hybridization signals appearing in all the DNA samples tested. This suggests a particularly important role of the gene corresponding to clone 1c6 through evolution and further supports the idea of its functional significance to biological processes in tissues where it is expressed, in turn emphasizing its status as a potential disease target.

A more broad analysis of conservation of the 72 putative retinal genes in the 6q12-q16 interval was performed by accessing information on the species in which homologous genes have been identified, as catalogued in the HomoloGene resource linked to the UniGene entries for each of the EST clusters. The majority (39/72; 54%) of the EST clusters representing human genes share homology with sequences that have been identified in at least one other species, with most of those (21/39; 54%)

corresponding to members of conserved gene families found in five or more species to date (Figure 3.5). This suggests that many of the putative retinal and/or RPE genes present in the chromosome 6q region of interest are likely to perform a critical biological function common to a number of different species, thereby further validating their categorization as candidate genes for human disease.

3.4 DISCUSSION

Both genetic sequence and expression data have been utilized to generate an integrated physical and transcript map for a region on human chromosome 6q known to coincide with multiple retinal disease loci. By focussing on sequences expressed in tissues directly relevant to retinal disease, this map is tailor-made for identifying genes that are important for the proper functioning of these tissues. In this way, specific sequences can be chosen over others within a given genomic interval of interest in the search for disease-causing mutations.

The map presented here represents a compilation of publicly available genome sequence information describing the nature and location of various genetic markers along the length of the human chromosome 6q12-q16 interval studied and the tissue-selective expression status of ESTs corresponding to both known as well as currently uncharacterized genes as deduced through hybridization experiments. As the current challenge of the Human Genome Project is the annotation of the entire genomic sequence that has been revealed, creation of such a map is an example of how additional information can be generated and integrated to contribute to an understanding of the significance of specific sequences.

Based on the analysis performed, all of the non-overlapping ESTs correctly mapping to and assessed within the D6S1596 – D6S1644 chromosome 6q interval were found to be expressed within retina-associated tissues. This appears to represent a fairly large number of retinal genes localizing to a relatively small segment of the human genome. Given the density of retinal diseases mapping to this genomic interval, perhaps it is reasonable to expect that such a region would be particularly rich in retinal genes if it is assumed that the different disorders are not attributable solely to allelic heterogeneity, environmental inflences, and the effects of genetic modifier loci. Thus if a particular genomic region is found to contain a high density of tissue-selective ESTs, it might be reasonable to predict that such a region may be predisposed towards genetic diseases that affect that particular tissue. This can then suggest specific genetic intervals and markers that can be tested preferentially for positive linkage in affected pedigrees rather than pursuing a genome-wide scan to pinpoint the position of a disease locus. Such associations between the concentration of retinal genes on particular chromosomes and presumed hotspots for retinal dysfunction have been previously observed.[34,37,38]

Aside from considering the overall retinal/RPE expression of the ESTs on the map, a striking demonstration of the utility of the approach adopted lies in the high proportion of genes uncovered that were not previously described to be but are now classified as retina or RPE-expressed. This information is of equivalent significance to both novel genes as well as those whose identity is known. Thus while some of the originally uncharacterized ESTs on chromosome 6q have since been found to correspond to particular genes of known identity, the finding that a number of these known genes are retina- or RPE-expressed genes is novel. This makes such genes, which would have otherwise been discounted based on their *in silico* expression profiles, relevant to the study of retinal function, and thereby enlarges the pool of genes that may be considered as candidates for retinal diseases mapping to chromosome 6q.

Prior report of the construction of a comprehensive high-resolution genomic transcript map of the human retina involved only an analysis of UniGene entries in which retina library clones were represented, and of those, only the entries for which a chromosomal map location was known were incorporated into the map.[37] As a result, 2 independent UniGene entries were localized to the 6q region corresponding to the D6S1596 – D6S1644 interval of interest, which sharply contrasts with the present finding of 60 such entries representing independent genes localized to this region. While some of this discrepancy can likely be accounted for by the updates that have recently been made to the human genome sequence and the growth of EST databases, it is evident that all genes that lie within a given genomic interval should be considered as potential candidates for expression within a given tissue until they can be definitively excluded rather than relying exclusively on *in silico* expression profiles given in gene indexing databases as accurate representations of *in vivo* expression status. A comprehensive transcript map must therefore incorporate both an exhaustive list of genes that have been localized to the genetic interval of interest and accurate experimental data pertaining to their expression within the particular tissue under study.

Towards further characterization of individual retina- or RPE-expressed sequences on our map, the expression patterns of the ESTs were elucidated both in sub-ocular and non-ocular human tissues. Despite the finding that none of the ESTs examined showed retina- or ocular-specific expression, the presence of these transcripts within retinal or RPE tissues at all classifies them as putative retinal disease candidates, as it is well established that tissue-specific expression is not necessarily a characteristic of genes giving rise to diseases that primarily affect a particular tissue, including diseases of the retina.[39]

Evolutionary conservation is another feature of genes that are thought to be important for biological function and therefore likely targets for dysfunction. Through examination of sequences of interest for their representation in the genomes of diverse species, a measure of conservation through evolution is enabled, which can serve as yet another criterion for prioritizing sequences for further investigation. The majority of the putative retinal and/or RPE genes on 6q12-q16 appear to be represented in the genomes of multiple species, lending further support to their classification as candidate disease genes. While a fair proportion (33/72; 46%) of the genes have only been identified in the human genome to date, many of these may not in fact correspond to human-specific genes but rather may have as yet unidentified homologues in species whose genomes are currently undergoing annotation as well as in those organisms for which genome sequencing is at present incomplete. The identification of non-human species in which gene homologues are found is in addition useful for suggesting possible model organisms in which the function of such genes and the effects of mutations in those genes can be studied.

In summary, human genome sequence data and bioinformatic tools developed to analyze genetic information have been used to generate a map that combines known positions of genetic markers with retinal and RPE expression status of ESTs corresponding to both known and as yet uncharacterized genes on a segment of human chromosome 6q. This in turn can serve as a resource in the search for candidate disease genes for retinal disorders mapping to 6q. It also represents a model upon which similar maps can be developed for other regions of the genome and for other tissue types, ultimately contributing to functional genomic studies that aim to decipher human genomic sequences.

3.5 ACKNOWLEDGEMENTS

The families and their clinicians are thanked for their contributions to these studies. This research was supported by the Canadian Institutes of Health Research (PSL), the E.A. Baker Foundation (PWW), the Canadian National Institute for the Blind (PSL, PWW), the Alberta Heritage Foundation for Medical Research (PSL, PWW), the Foundation Fighting Blindness-Canada (PWW), the Foundation Fighting Blindness USA (RA), the National Institutes of Health (research grant EY13198 and core grant EY07003) (RA), and a Research to Prevent Blindness Foundation Harrington Scholar Award (RA).

3.6 REFERENCES

1. Stone EM, Nichols BE, Kimura AE, Weingeist TA, Drack A, Sheffield VC. (1994) Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. *Arch. Ophthalmol.* 112:765-772.

2. Edwards AO, Miedziak A, Vrabec T, Verhoeven J, Acott TS, Weleber RG, Donoso LA. (1999) Autosomal dominant Stargardt-like macular dystrophy: I. Clinical characterization, longitudinal follow-up, and evidence for a common ancestry in families linked to chromosome 6q14. *Am. J. Ophthalmol.* 127:426-435.

3. Lagali PS, MacDonald IM, Griesinger IB, Chambers ML, Ayyagari R, Wong PW. (2000) Autosomal dominant Stargardt-like macular dystrophy segregating in a large Canadian family. *Can. J. Ophthalmol.* 35:315-324.

4. Griesinger IB, Sieving PA, Ayyagari R. (2000) Autosomal dominant macular atrophy at 6q14 excludes CORD7 and MCDR1/PBCRA loci. *Invest. Ophthalmol. Vis. Sci.* 41:248-255.

5. Ruiz A, Borrego S, Marcos I, Antinolo G. (1998) A major locus for autosomal recessive retinitis pigmentosa on 6q, determined by homozygosity mapping of chromosomal regions that contain gamma-aminobutyric acid-receptor clusters. *Am. J. Hum. Genet.* 62:1452-1459.

6. Khaliq S, Hameed A, Ismail M, Mehdi SQ, Bessant DA, Payne AM, Bhattacharya SS. (1999) Refinement of the locus for autosomal recessive Retinitis pigmentosa (RP25) linked to chromosome 6q in a family of Pakistani origin. *Am. J. Hum. Genet.* 65:571-574.

7. Kelsell RE, Gregory-Evans K, Gregory-Evans CY, Holder GE, Jay MR, Weber BH, Moore AT, Bird AC, Hunt DM. (1998) Localization of a gene (CORD7) for a dominant cone-rod dystrophy to chromosome 6q. *Am. J. Hum. Genet.* 63:274-279.

8. Small KW, Udar N, Yelchits S, Klein R, Garcia C, Gallardo G, Puech B, Puech V, Saperstein D, Lim J, Haller J, Flaxel C, Kelsell R, Hunt D, Evans K, Lennon F, Pericak-Vance M. (1999) North Carolina macular dystrophy (MCDR1) locus: a fine resolution genetic map and haplotype analysis. *Mol. Vis.* 5:38.

9. Kelsell RE, Godley BF, Evans K, Tiffin PA, Gregory CY, Plant C, Moore AT, Bird AC, Hunt DM. (1995) Localization of the gene for progressive bifocal chorioretinal atrophy (PBCRA) to chromosome 6q. *Hum. Mol. Genet.* 4:1653-1656.

10. Dharmaraj S, Li Y, Robitaille JM, Silva E, Zhu D, Mitchell TN, Maltby LP, Baffoe-Bonnie AB, Maumenee IH. (2000) A novel locus for Leber congenital amaurosis maps to chromosome 6q. *Am. J. Hum. Genet.* 66:319-326.

11. Kniazeva M, Traboulsi EI, Yu Z, Stefko ST, Gorin MB, Shugart YY, O'Connell JR, Blaschak CJ, Cutting G, Han M, Zhang K. (2000) A new locus for dominant drusen and macular degeneration maps to chromosome 6q14. *Am. J. Ophthalmol.* 130:197-202.

12. van Lith-Verhoeven JJ, Hoyng CB, van den Helm B, Deutman AF, Brink HM, Kemperman MH, de Jong WH, Kremer H, Cremers FP. (2004) The benign concentric annular macular dystrophy locus maps to 6p12.3-q16. *Invest. Ophthalmol. Vis. Sci.* 45:30-5.

13. Rivas F, Ruiz C, Rivera H, Moller M, Serrano-Lucas JI, Cantu JM. (1986) De novo del(6)(q25) associated with macular degeneration. *Ann. Genet.* 29:42-44.

14. Pierpont ME, MacCarthy KG, Knobloch WH. (1986) Partial trisomy 6q and bilateral retinal detachment. *Ophthalmic Paediatr. Genet.* 7:175-180.

15. Tranebjaerg L, Sjo O, Warburg M. (1986) Retinal cone dysfunction and mental retardation associated with a de novo balanced translocation 1;6(q44;q27). *Ophthalmic Paediatr. Genet.* 7:167-173.

16. McLeod DR, Fowlow SB, Robertson A, Samcoe D, Burgess I, Hoo JJ. (1990) Chromosome 6q deletions: a report of two additional cases and a review of the literature. *Am. J. Med. Genet.* 35:79-84.

17. Warburg M, Tranebjaerg L, Sjo O. (1990) Retinal cone dystrophy and chromosome 6(q26) deletions. *Ann. Genet.* 33:124.

18. Warburg M, Sjo O, Tranebjaerg L. (1991) Chromosome 6q deletion and retinal cone dystrophy. *Am. J. Med. Genet.* 38:134.

19. Zhang K, Kniazeva M, Han M, Li W, Yu Z, Yang Z, Li Y, Metzker ML, Allikmets R, Zack DJ, Kakuk LE, Lagali PS, Wong PW, MacDonald IM, Sieving PA, Figueroa DJ, Austin CP, Gould RJ, Ayyagari R, Petrukhin K. (2001) A 5-bp deletion in *ELOVL4* is associated with two related forms of autosomal dominant macular dystrophy. *Nat. Genet.* 27:89-93.

20. Lagali PS, Ayyagari R, Wong PW. (2004) Updating the candidate gene approach to study human macular degenerative disease: applied bioinformatics. *TIBETS* (in press).

21. Johnson S, Halford S, Morris AG, Patel RJ, Wilkie SE, Hardcastle AJ, Moore AT, Zhang K, Hunt DM. (2003) Genomic organisation and alternative splicing of human *RIM1*, a gene implicated in autosomal dominant cone-rod dystrophy (CORD7). *Genomics* 81:304-314.

22. van den Brink DM, Brites P, Haasjes J, Wierzbicki AS, Mitchell J, Lambert-Hamill M, de Belleroche J, Jansen GA, Waterham HR, Wanders RJ. (2003) Identification of *PEX7* as the second gene involved in Refsum disease. *Am. J. Hum. Genet.* 72:471-477.

23. Muntau AC, Mayerhofer PU, Paton BC, Kammerer S, Roscher AA. (2000) Defective peroxisome membrane synthesis due to mutations in human *PEX3* causes Zellweger syndrome, complementation group G. *Am. J. Hum. Genet.* 67:967-975.

24. Ghaedi K, Honsho M, Shimozawa N, Suzuki Y, Kondo N, Fujiki Y. (2000) *PEX3* is the causal gene responsible for peroxisome membrane assembly-defective Zellweger syndrome of complementation group G. *Am. J. Hum. Genet.* 67:976-981.

25. Stanesu-Segal B, Evrard P. (1989) Zellweger syndrome, retinal involvement. *Metab. Pediatr. Syst. Ophthalmol.* 12:96-99.

26. Felbor U, Gehrig A, Sauer CG, Marquardt A, Kohler M, Schmid M, Weber BH. (1998) Genomic organization and chromosomal localization of the interphotoreceptor matrix proteoglycan-1 (*IMPG1*) gene: a candidate for 6q-linked retinopathies. *Cytogenet. Cell. Genet.* 81:12-17.

27. Cutting GR, Curristin S, Zoghbi H, O'Hara B, Seldin MF, Uhl GR. (1992) Identification of a putative gamma-aminobutyric acid (GABA) receptor subunit rho2 cDNA and colocalization of the genes encoding rho2 (*GABRR2*) and rho1 (*GABRR1*) to human chromosome 6q14-q21 and mouse chromosome 4. *Genomics* 12:801-806.

28. Bailey ME, Albrecht BE, Johnson KJ, Darlison MG. (1999) Genetic linkage and radiation hybrid mapping of three human GABA(C) receptor rho subunit genes: *GABRR1*, *GABRR2* and *GABRR3*. *Biochim. Biophys. Acta.* 1447:307-312.

29. Yarovinsky TO, Rickman DW, Diamond RH, Taub R, Hageman GS, Bowes Rickman C. (2000) Expression of the protein tyrosine phosphatase, phosphatase of regenerating liver 1, in the outer segments of primate cone photoreceptors. *Brain Res. Mol. Brain Res.* 77:95-103.

30. Lagali PS, Kakuk LE, Griesinger IB, Wong PW, Ayyagari R. (2002) Identification and Characterization of *C6orf37*, a Novel Candidate Human Retinal Disease Gene on Chromosome 6q14. *Biochem. Biophys. Res. Commun.* 293(1):356-365.

31. Gehrig A, Felbor U, Kelsell RE, Hunt DM, Maumenee IH, Weber BH. (1998) Assessment of the interphotoreceptor matrix proteoglycan-1 (*IMPG1*) gene localised

to 6q13-q15 in autosomal dominant Stargardt-like disease (ADSTGD), progressive bifocal chorioretinal atrophy (PBCRA), and North Carolina macular dystrophy (MCDR1). *J. Med. Genet.* 35:641-645.

32. Kniazeva MF, Chiang MF, Cutting GR, Zack DJ, Han M, Zhang K. (1999) Clinical and genetic studies of an autosomal dominant cone-rod dystrophy with features of Stargardt disease. *Ophthalmic Genet.* 20:71-81.

33. Marcos I, Ruiz A, Blaschak CJ, Borrego S, Cutting GR, Antinolo G. (2000) Mutation analysis of *GABRR1* and *GABRR2* in autosomal recessive retinitis pigmentosa. *J. Med. Genet.* 37:E5.

34. Blackshaw S, Fraioli RE, Furukawa T, Cepko CL. (2001) Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. *Cell* 107:579-589.

35. Mazzocco M, Maffei M, Egeo A, Vergano A, Arrigo P, Di Lisi R, Ghiotto F, Scartezzini P. (2002) The identification of a novel human homologue of the SH3 binding glutamic acid-rich (*SH3BGR*) gene establishes a new family of highly conserved small proteins related to Thioredoxin Superfamily. *Gene* 291:233-239.

36. Gonzalez-Manchon C, Ferrer M, Ayuso MS, Parrilla R. (1995) Cloning, sequencing and functional expression of a cDNA encoding a NADP-dependent malic enzyme from human liver. *Gene* 159:255-260.

37. Bortoluzzi S, d'Alessi F, Danieli GA. (2000) A novel resource for the study of genes expressed in the adult human retina. *Invest. Ophthalmol. Vis. Sci.* 2000 41:3305-3308.

38. Yu J, Farjo R, MacNee SP, Baehr W, Stambolian DE, Swaroop A. (2003) Annotation and analysis of 10,000 expressed sequence tags from developing mouse eye and adult retina. *Genome Biol.* 4:R65.

39. Pacione LR, Szego MJ, Ikeda S, Nishina PM, McInnes RR. (2003) Progress toward understanding the genetic and biochemical mechanisms of inherited photoreceptor degenerations. *Annu. Rev. Neurosci.* 26:657-700.

Figure 3.1 Dot blot analysis of chromosome 6q12-q16 ESTs for determining retinal/RPE-expression status. A sample representation of chromosome 6q EST clones arrayed onto nylon membranes and probed with human peripheral retina-, fovea-, or retinal pigment epithelium (RPE)-derived total cDNA probes is shown. Control dot blots using tissue-specific or tissue-predominant markers are displayed on the right. See text for explanation of relative hybridization level scoring and differential expression observed for representative ESTs.

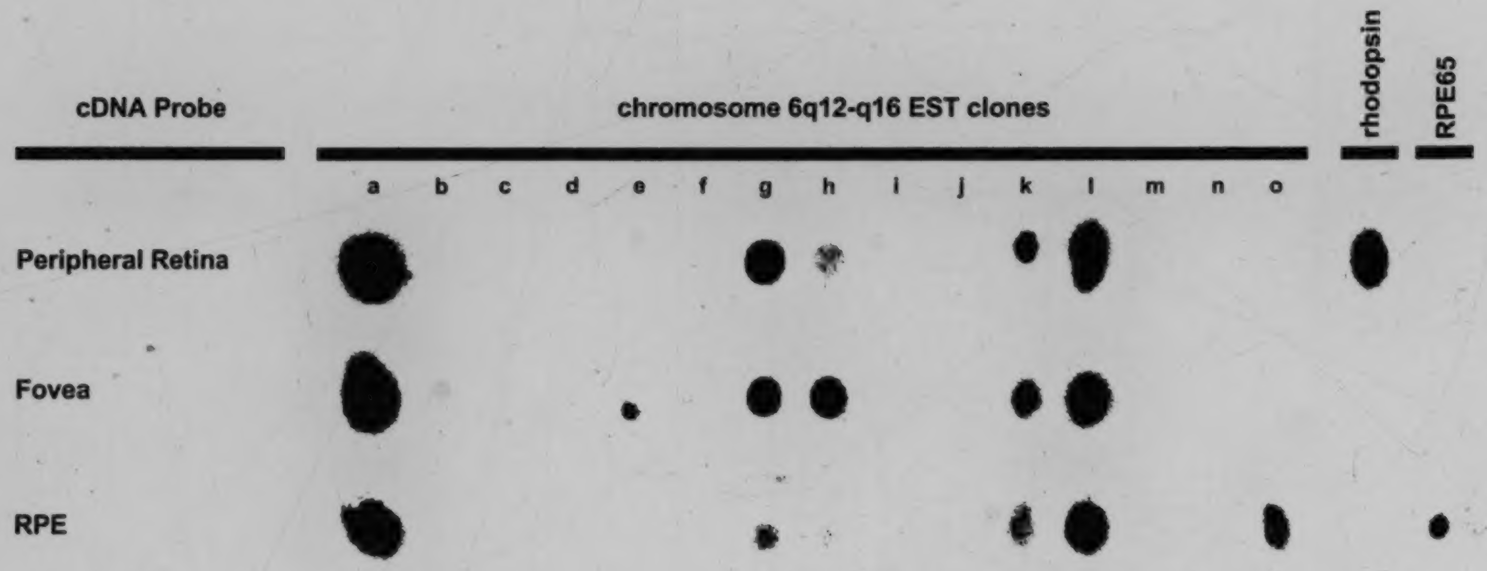

cDNA Probe
chromosome 6q12-q16 EST clones
rhodopsin
RPE65
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
Peripheral Retina
Fovea
RPE

Figure 3.2 Proportion and classification of genes represented by chromosome 6q12-q16 EST clones hybridizing to human retina/RPE-derived cDNA probes. Hatched bars correspond to genes previously known to be expressed in the tissue indicated below the graph based on derivative libraries of constituent ESTs as reported in the UniGene database. Stippled bars correspond to retina and/or RPE-expressed genes based solely on this study. Solid bars correspond to genes found to be enriched or expressed in the indicated tissue type relative to the other tissues examined in this study.

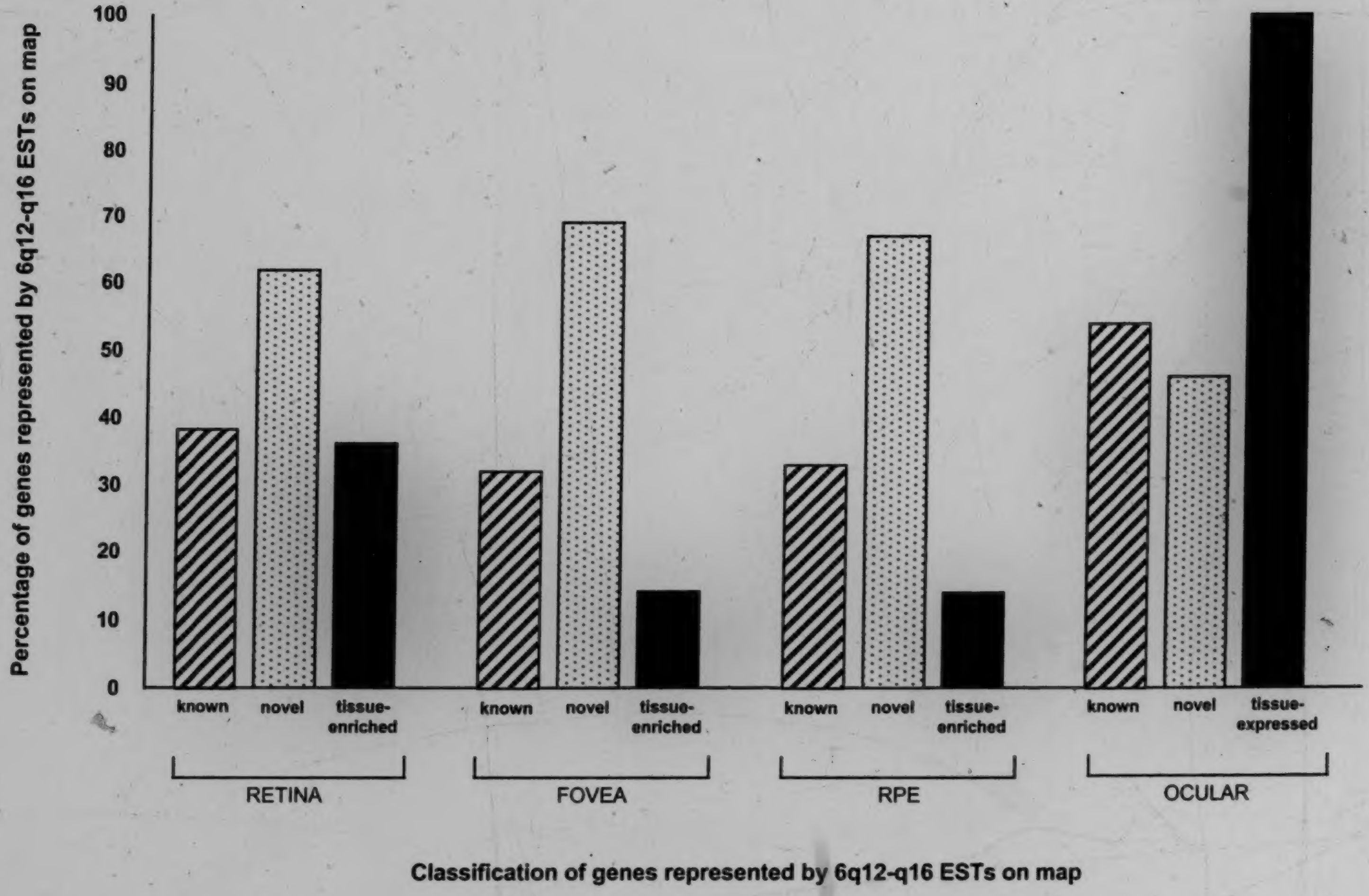

Percentage of genes represented by 6q12-q16 ESTs on map
100
90
80
70
60
50
40
30
20
10
0
known
novel
tissue-enriched
RETINA
known
novel
tissue-enriched
FOVEA
known
novel
tissue-enriched
RPE
known
novel
tissue-expressed
OCULAR
Classification of genes represented by 6q12-q16 ESTs on map

Table 3.1 A retinal and RPE transcript map for a segment of human chromosome 6q. Microsatellite markers (M), known genes (G), predicted genes (P), pseudogenes (S), and ESTs (E) are listed in order according to reported nucleotide positions on chromosome 6. Numerical values representing relative expression levels of human ESTs in peripheral retina (Per.Ret.), fovea, and RPE tissue are based on hybridization intensities observed for macroarrays probed with tissue-derived total cDNA pools (see text for details); average hybridization levels are given for map entries corresponding to loci comprised of multiple ESTs. The 6q12-q16 chromosomal interval represented on the map between microsatellite markers D6S1596 and D6S1644 contains a more comprehensive list of genetic elements than the flanking regions shown. *na*, not available.

Entry	6q band	Nucleotide Position	Marker	Uni(Gene/STS) ID	GenBank Accession
M		71514182-71514287	D6S455	UniSTS: 11491	na
G	6q13	71566964-71661730	B3GAT2	UniGene: Hs.4273	NM_080742
E	6q12-q13	71567146-71567272	EST 3g7	UniGene: Hs.4273	AI536849
S	6q13	71870342-71872761	LOC157713	na	NG_001063
G		71994109-72007555	FLJ21079	UniGene: Hs.16512	NM_024576.2
G	6q13	72119747-72126046	C6orf155	UniGene: Hs.368337	NM_024882
S	6q13	72290022-72291209	LOC377402	na	na
G	6q12-q13	72592004-72969301	RIMS1	UniGene: Hs.302136	AY190519
P	6q14	73303777-73304960	LOC340167	UniGene: Hs.485730	XM_294054
G	6q14	73327433-73901178	KCNQ5	UniGene: Hs.371688	AF202977
M		73405576-73405722	D6S1036	UniSTS: 39999	G09175
M		73741270-73741427	D6S280	UniSTS: 32427	Z16764
G	6q14.1	73914191-73968517	C6orf148	UniGene: Hs.412173	NM_030568
S	6q14.1	73997263-73998730	EIF3S6P1	na	NG_001023
E	6q14.1	73998587-73998721	EST 1b8	UniGene: Hs.412173	R55826
S	6q14.1	74028167-74028832	LOC377403	na	na
S	6q14.1	74030676-74034338	LOC377404	na	na
P	6q14.1	74058385-74059551	LOC340168	UniGene: Hs.125331	XM_291161
E		74058511-74058672	EST 3g2	UniGene: Hs.125331	AI365263
P	6q14.1	74067998-74069181	LOC154288	UniGene: Hs.447749	XM_087901
P	6q14.1	74073880-74075113	LOC375517	na	XM_351665
G	6q12-q13	74100069-74122668	DDX43	UniGene: Hs.125507	NM_018665
G	6q14.1	74130438-74157597	C6orf150	UniGene: Hs.344080	NM_138441
G	6q14.1	74167086-74206773	MTO1	UniGene: Hs.33979	NM_012123
G	6q14.1	74222839-74226352	EEF1A1	UniGene: Hs.439552	NM_001402
G	6q14-q15	74298701-74359297	SLC17A5	UniGene: Hs.117865	NM_012434
M	6q14.1	74386624-74386834	D6S1596	UniSTS: 43426	Z52907
G	6q14.1	74401425-74530386	CD109	UniGene: Hs.94030	AF410459
M	6q14.1	74466307-74466510	D6S406	UniSTS: 33615	Z23304
E	6q14.1	74533336-74533470	EST 3e6	UniGene: Hs.94030	AI478383
M		74793692-74793901	D6S1622	UniSTS: 58591	Z53310
E		74974584-74974712	EST 3c5	UniGene: Hs.55246	W02146
M		75104984-75105174	D6S1659	UniSTS: 13962	Z53823
E		75395802-75395942	EST 3c6	UniGene: Hs.384600	W73409
G	6q12-q13	75789640-75911167	COL12A1	UniGene: Hs.101302	AJ420501,U73778
E		75789654-75789783	ESTs 1h11, 3e4	UniGene: Hs.101302	AA663291
E		75844846-75844977	EST 1g5	UniGene: Hs.71439	AA131469
G	6q12	75943100-75949122	COX7A2	UniGene: Hs.70312	AF134406, X15822
E		75958302-75959303	ESTs 3f4, 2b8, 3g9	UniGene: Hs.108530	AI191099, AI253345, AI597778
G	6q14.1	75958928-75990098	C6orf67	UniGene: Hs.108530	NM_018247
G	6q14.1	76013405-76120291	FILIP1	UniGene: Hs.102796	XM_029179
M		76079387-76079527	D6S2140	UniSTS: 20109	G13827
M		76102723-76102968	D6S456	UniSTS: 47860	Z24265
M		76303106-76303310	D6S990	UniSTS: 73512	G02945
G	6q13-q14.3	76307838-76421778	SUSP1	UniGene: Hs.435628	AF196304
E		76385034-76421922	ESTs 2a7, 3a6, 2a11, 3g10, 3a10, 1c2	UniGene: Hs.435628	AA768785, R05820, AA910527, AI609926, R66010, R27598
M		76440124-76440265	D6S1999	UniSTS: 43249	G14151
G	6q13	76454646-76621541	MYO6	UniGene: Hs.118483	AB002387
M		76593945-76594125	D6S1986	UniSTS: 29766	G14049
E		76598321-76621964	ESTs 1g4, 3d3, 3c9	UniGene: Hs.118483	AA030004, AA129322, AA025570
G	6q14.2-q15	76626659-76777932	IMPG1	UniGene: Hs.129882	AF047492
E		76656168-76709266	EST 3b4	UniGene: Hs.129882	H38604
M		77827558-77827720	D6S1625	UniSTS: 46010	Z53335
M		78100659-78100901	AFM207YE7	UniSTS: 13277	Z66785
G	6q13	78167544-78168716	HTR1B	UniGene: Hs.520239	NM_000863
E		78169926-78170058	EST 3d2	UniGene: Hs.68065	AA083643
M		78451904-78452085	D6S1589	UniSTS: 32002	Z52861
E		79179158-79179307	EST 1a8	na	T89351

Description	Dot Blot Expression		
	Per.Ret.	Fovea	RPE
microsatellite marker			
gene; beta-1,3-glucuronyltransferase 2 (glucuronosyltransferase S)			
6qEST; B3GAT2	2	2	4
pseudogene; lysophospholipase I-like			
gene; hypothetical protein FLJ21079			
gene; chromosome 6 open reading frame 155; hypothetical protein FLJ13189			
pseudogene; LOC377402: similar to keratin 19; keratin, type I cytoskeletal 19; keratin, type I, 40 kd; cytokeratin 19; 40-kDa keratin intermediate filament precursor gene			
gene; regulating synaptic membrane exocytosis 1 (CORD7 gene)			
predicted gene; similar to hypothetical protein DKFZp434A1315			
gene; potassium voltage-gated channel, KQT-like subfamily, member 5			
microsatellite marker			
microsatellite marker			
gene; chromosome 6 open reading frame 148			
pseudogene; eukaryotic translation initiation factor 3, subunit 6 48kDa pseudogene 1			
6qEST; C6orf148	4	5	2
pseudogene; LOC377403: similar to syntenin-2alpha; syntenin-2			
pseudogene; LOC377404: similar to multifunctional protein ADE2			
predicted gene; hypothetical protein similar to developmental pluripotency associated 5; embryonal stem cell specific gene 1			
6qEST; similar to developmental pluripotency associated 5; embryonal stem cell specific gene 1	2	2	3
predicted gene; hypothetical protein			
predicted gene; similar to RIKEN cDNA 2410146L05; hypothetical protein			
gene; DEAD (Asp-Glu-Ala-Asp) box polypeptide 43			
gene; chromosome 6 open reading frame 150; hypothetical protein BC012928			
gene; mitochondrial translation optimization 1 homolog of S. cerevisiae Mto1			
gene; eukaryotic translation elongation factor 1 alpha 1			
gene; solute carrier family 17 (anion/sugar transporter), member 5			
microsatellite marker			
gene; CD109 cell surface antigen (Gov platelet alloantigens)			
microsatellite marker			
6qEST; CD109 antigen (Gov platelet alloantigens)	2	3	3
microsatellite marker			
6qEST; Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone	3	3	3
microsatellite marker			
6qEST; full length insert cDNA clone ZD50H02; Soares_fetal_heart_NbHH19W cDNA clone	3	3	3
gene; collagen, type XII, alpha 1			
6qESTs; COL12A1	3	3	3
6qEST; Soares_pregnant_uterus_NbHPU cDNA clone	2	1	1
gene; cytochrome c oxidase subunit VIIa polypeptide 2 (liver)			
6qESTs; C6orf67	2	3	3
gene; chromosome 6 open reading frame 67; hypothetical protein FLJ10856			
gene; filamin A interacting protein 1			
microsatellite marker			
microsatellite marker			
microsatellite marker			
gene; SUMO-1-specific protease			
6qESTs; SUSP1	3	4	3
microsatellite marker			
gene; myosin VI (DFNA22, DFNB37 gene)			
microsatellite marker			
6qESTs; MYO6	2	4	3
gene; interphotoreceptor matrix proteoglycan 1			
6qEST; IMPG1	5	5	3
microsatellite marker			
microsatellite marker			
gene; 5-hydroxytryptamine (serotonin) receptor 1B			
6qEST; Stratagene endothelial cell 937223 cDNA clone	1	4	4
microsatellite marker			
6qEST; Soares fetal liver spleen 1NFLS cDNA clone	5	4	4

Entry	6q band	Nucleotide Position	Marker	Uni(Gene/STS) ID	GenBank Accession
M		79340882-79341128	D6S284	UniSTS: 28344	Z16822
M		79340908-79341019	AFMA121YH5	UniSTS: 60690	Z67307
G	6q14-q15	79572867-79604016	IRAK1BP1	UniGene: Hs.511817	XM_059729, AL049321
M		79609643-79609844	D6S1481	UniSTS: 76514	G04675
E		79641236-79646058	ESTs 3e10, 3b3, 1g12	UniGene: Hs.9927	AI091194, H30653, AA167774
G	6q14	79645862-79675437	PHIP	UniGene: Hs.9927	NM_017934
G	6q14.3	79906560-79940051	HMGN3	UniGene: Hs.77558	AY043282, BI551210
E		79906583-79942061	ESTs 1h4, 2a3, 3d10	UniGene: Hs.77558	AA215656, AA632557, AA256476
M		80039742-80039944	D6S1065	UniSTS: 16594	na
S	6q14.3	80111588-80123612	LOC377405	na	na
P	6q14.3	80125610-80126972	LOC153959	UniGene: Hs.435465	XM_098450
M		80128068-80128217	D6S1465	UniSTS: 38985	G04492
S	6q14.3	80141993-80142539	DBIL2	na	NG_001112
G	6q14.3	80190305-80242743	C6orf152	UniGene: Hs.21945	NM_181714.1
E		80190383-80192077	ESTs 1b4, 3a8	UniGene: Hs.21945	R42841, R21456
E		80290179-80290363	ESTs 1f3, 3e7	UniGene: Hs.55607	W37776, AA854658
G	6q13-q15	80336596-80408965	SH3BGRL2	UniGene: Hs.302772	AF340151
M		80346465-80346838	D6S460	UniSTS: 6530	Z24323
E		80407590-80408898	ESTs 1d1, 1f6, 1e1	UniGene: Hs.302772	H19036, AA019723, H57896
G		80507728-80511304	C6orf7	UniGene: Hs.288213	AI631615, AA774200
G	6q14	80620125-80652818	ELOVL4	UniGene: Hs.101915	AY037298
E		80620155-80621337	ESTs 2b2, 3e11	UniGene: Hs.101915	AA993222, AA969661
S	6q14.3	80658318-80659324	LOC377406	na	na
G	6q13-q21	80709953-80747833	TTK	UniGene: Hs.169840	NM_003318
M		80743134-80743290	D6S2407	UniSTS: 29430	G10387
P	6q14.3	80769598-80775326	LOC285786	UniGene: Hs.446411	XM_208349
G	6q13-q15	80811960-81051583	BCKDHB	UniGene: Hs.1265	M55575
E		80950465-80950626	EST 3e8	UniGene: Hs.128606	AA907027
E		80986913-81051552	EST 3a3, 3d9	UniGene: Hs.1265	T59946, AA262255
M		80986981-80987117	D6S1856	UniSTS: 80919	G11507, T59946
M		81049159-81049259	D6S1762	UniSTS: 32076	G07116, M55575
E		81073642-81073817	EST 1e3	UniGene: Hs.418518	H60737
P	6q14.3	81079328-81079882	LOC377407	na	XM_352618
P	6q14.3	81148703-81168520	LOC340171	UniGene: Hs.386402	XM_295178
E		81168116-81168487	ESTs 3c11, 1e5, 3f8	UniGene: Hs.386402	AA058574, H79994, AI273723
P	6q14.3	81261154-81262502	LOC153956	UniGene: Hs.485835	XM_094621
M		81268081-81268346	D6S1707	UniSTS: 4631	Z51846
M		81883004-81883134	D6S463	UniSTS: 29333	Z24349
M		82000473-82000623	D6S251	UniSTS: 18071	L00999
M		82190613-82190881	D6S1646	UniSTS: 6759	Z51201
E		82222301-82222426	EST 1c4	UniGene: Hs.390240	R62950
E		82227889-82228070	EST 1d11	na	H54787
M		82442348-82442601	D6S1455	UniSTS: 60319	G04378
M		82450368-82450594	D6S445	UniSTS: 70247	Z24053
E		82451108-82457437	ESTs 1e7, 1h10, 3b10, 1b12, 1e4, 1e6	UniGene: Hs.10784	H92739, AA460873, H97670, R28669, H65844, H92608
M		82453007-82453110	D6S1495	UniSTS: 37750	G04782
G	6q14	82454955-82457992	C6orf37	UniGene: Hs.10784	AF350451
P	6q14.3	82550853-82643668	LOC375518	na	XM_353402
G	6q14.3	82875552-82897040	IBTK	UniGene: Hs.306425	AF235049
M	6q14.3	82875587-82875690	D6S1882	UniSTS: 39699	G11689, T50208
E		82875587-82875690	ESTs 3e5, 1h5	UniGene: Hs.306425	AA748404, AA236615
M	6q14.3	82942817-82943130	D6S1780	UniSTS: 1904	G10001
M		83003767-83004148	D6S954	UniSTS: 30928	L17875
G	6q14-q15	83069135-83072210	TPBG	UniGene: Hs.82128	Z29083
E		83075899-83076014	EST 1g6	UniGene: Hs.71947	AA149121
E		83478639-83478757	EST 1c9	UniGene: Hs.485848	H06999
G	6q15	83597785-83771146	C6orf157	UniGene: Hs.406722	AL137544
M		83606562-83606787	D6S1458	UniSTS: 20753	G04426

Description	Dot Blot Expression		
	Per.Ret.	Fovea	RPE
microsatellite marker			
microsatellite marker			
gene; Interleukin-1 receptor-associated kinase 1 binding protein 1			
microsatellite marker			
6qESTs; PHIP	1	3	3
gene; pleckstrin homology domain interacting protein			
gene; high mobility group nucleosomal binding domain 3			
6qESTs; HMGN3	3	3	2
microsatellite marker			
pseudogene; LOC377405: similar to Erythrocyte band 7 integral membrane protein (Stomatin)			
predicted gene; LOC153959 hypothetical protein			
microsatellite marker			
pseudogene; diazepam binding inhibitor-like 2			
gene; chromosome 6 open reading frame 152			
6qESTs; C6orf152	2	3	3
6qESTs; Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clones	5	5	3
gene; SH3 domain binding glutamic acid-rich protein like 2			
microsatellite marker			
6qESTs; SH3BGRL2	4	4	2
gene; chromosome 6 open reading frame 7			
gene; elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 4			
6qESTs; ELOVL4	3	3	3
pseudogene; similar to Glyceraldehyde 3-phosphate dehydrogenase, liver (GAPDH)			
gene; TTK protein kinase			
microsatellite marker			
predicted gene; similar to KIAA1424 protein; hypothetical protein			
gene; branched chain keto acid dehydrogenase E1, beta polypeptide			
6qEST; Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone	1	3	3
6qEST; BCKDHB	3	4	5
microsatellite marker			
microsatellite marker			
6qEST; Soares fetal liver spleen 1NFLS cDNA clone	5	4	3
predicted gene; similar to Ac2-210; hypothetical protein			
predicted gene; hypothetical protein			
6qESTs; LOC340171	3	4	3
predicted gene; similar to citrate synthase precursor; Citrate synthase, mitochondrial			
microsatellite marker			
microsatellite marker			
microsatellite marker			
microsatellite marker			
6qEST; Soares placenta Nb2HP cDNA clone	3	3	2
6qEST; Soares fetal liver spleen 1NFLS cDNA clone	4	4	2
microsatellite marker			
microsatellite marker			
6qESTs; C6orf37	3	3	4
microsatellite marker			
gene; chromosome 6 open reading frame 37; hypothetical protein FLJ20037			
predicted gene; hypothetical protein			
gene; inhibitor of Bruton agammaglobulinemia tyrosine kinase			
microsatellite marker			
6qESTs; IBTK	1	2	2
microsatellite marker			
microsatellite marker			
gene; trophoblast glycoprotein			
6qEST; Soares_pregnant_uterus_NbHPU cDNA clone	2	3	2
6qEST; Soares infant brain 1NIB cDNA clone	4	3	2
gene; chromosome 6 open reading frame 157; hypothetical protein DKFZp434A1520			
microsatellite marker			

Entry	6q band	Nucleotide Position	Marker	Uni(Gene/STS) ID	GenBank Accession
G	6q15	83843331-83873784	KIAA1117	UniGene; Hs.348982	NM_015018
E		83873685-83876535	ESTs 1c11, 3c2	UniGene: Hs.348982	H09058, N52341
G	6q14.1-q15	83874226-83898531	PGM3	UniGene: Hs.237323	AF180371
G	6q15	83898729-83901847	RWDD2	UniGene: Hs.423336	NM_033411
E		83901694-83904100	EST 3f6, 1h8	UniGene: Hs.423336	AI219946, AA418452
E		83915733-83917195	ESTs 1a11, 3b12, 1e2, 3b7, 1b10, 3f7	UniGene: Hs.14732	R39955, N35825, H58541, R60023, AI339685
G	6q12	83916853-84136376	ME1	UniGene: Hs.14732	NM_002395
M		83942470-83942812	D6S391	UniSTS: 10427	L16349
M		83974000-83974209	D6S1020	UniSTS: 57829	G08491
E		84046233-84046461	EST 3c12	UniGene: Hs.21160	AI076834
G	6q15	84217867-84231015	C6orf158	UniGene: Hs.98381	BC037170
E		84230800-84230926	EST 3g5	UniGene: Hs.98381	AI424866
G	6q15	84258209-84414372	SNAP91	UniGene: Hs.304330	AB014556
E		84258402-84258520	EST 3f12	UniGene: Hs.304330	AI341689
M		84381113-84381284	D6S1609	UniSTS: 62917	Z53127
M		84386308-84386509	D6S1922	UniSTS: 11302	G04644
P	6q15	84558581-84562828	LOC134701	UniGene: Hs.149454	XM_059730
G	6pter-q22.33	84565001-84665551	NCB5OR	UniGene: Hs.5741	NM_016230
M		84624997-84625268	D6S1634	UniSTS: 65296	Z53468
E		84665524-84665697	ESTs 1h6, 3f11	UniGene: Hs.5741	AA287098, AI368056
G	6q15	84739016-84796195	LOC112609	UniGene: Hs.425033	NM_138409
E		84795905-84796175	ESTs 1b9, 3e2	UniGene: Hs.425033	R61804, AA524400
G	6q15	84829568-84926485	C6orf84	UniGene: Hs.202276	NM_014895
E		84830256-84830508	ESTs 3g6, 2a10	UniGene: Hs.202276	AI474053, AI147106
E		84832920-84833039	EST 1b7	UniGene: Hs.485863	R35585
M		85078224-85078423	D6S1252	UniSTS: 58937	G02936
P	6q15	85134895-85136115	LOC346033	UniGene: Hs.171384	XM_294042
M		85275673-85275912	D6S2416	UniSTS: 8249	G10403
M		85402923-85403030	D6S1627	UniSTS: 31945	Z53360
E		85416257-85416414	EST 3c7	na	AA063577
E		85437917-85438304	EST 3a12	UniGene: Hs.27705	H07146
G	6q14-q15	85441999-85469495	TBX18	UniGene: Hs.251830	XM_291165
M		85442926-85443027	D6S1475	UniSTS: 72211	G04546
M		85488089-85488363	AFMA119WD9	UniSTS: 48888	Z67303
M		85578961-85579079	D6S1045	UniSTS: 10841	G08518, G09620
M		85580233-85580375	D6S1601	UniSTS: 44162	Z51126
M		85704035-85704235	D6S1474	UniSTS: 4248	G04544
M		85869112-85869412	D6S1407	UniSTS: 26517	G04087
M		85950001-85950227	D6S1652	UniSTS: 41680	Z53713
S	6q15	85956702-85963002	LOC377408	na	na
P	6q15	85993040-85994305	LOC377409	UniGene: Hs.453465	XM_294043
S	6q15	86127188-86153094	LOC377410	na	na
G	6q14-q21	86155405-86201092	NT5E	UniGene: Hs.153952	X55740
E		86157236-86157389	EST 3e12	UniGene: Hs.153952	AI146540
G	6q15	86210813-86299213	SNX14	UniGene: Hs.512657	NM_153816, NM_020468
E		86210814-86211257	EST 1e11	UniGene: Hs.512657	N63097
E		86304398-86304572	EST 3d5	UniGene: Hs.485874	AA135142
E		86313252-86319915	ESTs 3d8, 1h7, 3a11, 3b1, 1d3	UniGene: Hs.436376	AA228115, AA479060, R75834, H20156
G	6q14-q15	86319535-86348623	SYNCRIP	UniGene: Hs.436376	NM_006372
E		86365175-86366394	ESTs 1a6, 3a4	UniGene: Hs.277418	T73987, T82459
E		86377937-86378118	EST 3c4	UniGene: Hs.245367	AI668667
S	6q15	86791742-86792469	LOC377411	na	na
P	6q15	86919546-86919899	LOC134686	UniGene: Hs.485885	XM_068985
M		87387753-87388001	D6S1291	UniSTS: 62160	G02949
M		87449108-87449356	AFMB354XG1	UniSTS: 29611	Z67686
M		87622138-87622367	AFM342XF9	UniSTS: 51453	Z67086
G	6q14-q15	87643006-87721992	HTR1E	UniGene: Hs.1611	M91467
G	6q12-q21	87790818-87800420	CGA	UniGene: Hs.119689	NM_000735

Description	Dot Blot Expression		
	Per.Ret.	Fovea	RPE
gene; KIAA1117 protein			
6qESTs; KIAA1117 protein	5	5	5
gene; phosphoglucomutase 3			
gene; RWD domain containing 2			
6qESTs; RWDD2	1	2	3
6qESTs; ME1	4	4	4
gene; malic enzyme 1, NADP(+)-dependent, cytosolic, malate dehydrogenase			
microsatellite marker			
microsatellite marker			
6qEST; Soares_fetal_liver_spleen_1NFLS_S1 cDNA clone	5	4	4
gene; chromosome 6 open reading frame 158 ; hypothetical protein MGC46520			
6qEST; C6orf158	1	3	2
gene; synaptosomal-associated protein, 91kDa homolog (mouse)			
6qEST; SNAP91	5	4	3
microsatellite marker			
microsatellite marker			
predicted gene; similar to Down syndrome critical region protein 6			
gene; NADPH cytochrome B5 oxidoreductase			
microsatellite marker			
6qESTs; NCB5OR	2	3	3
gene; chromosome 6 open reading frame 117			
6qESTs; C6orf117	4	3	3
gene; chromosome 6 open reading frame 84; KIAA1009 protein			
6qESTs; C6orf84	2	3	2
6qEST; Soares infant brain 1NIB cDNA clone	2	3	2
microsatellite marker			
predicted gene; similar to SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1			
microsatellite marker			
microsatellite marker			
6qEST; Soares_fetal_heart_NbHH19W cDNA clone	5	5	5
6qEST; Soares infant brain 1NIB cDNA clone	1	2	3
gene; T-box 18 transcription factor			
microsatellite marker			
microsatellite marker			
microsatellite marker			
microsatellite marker			
microsatellite marker			
microsatellite marker			
microsatellite marker			
pseudogene; similar to ribosomal protein L31			
predicted gene; similar to Keratin, type I cytoskeletal 18 (Cytokeratin 18)			
pseudogene; similar to Translationally controlled tumor protein (TCTP) (p23) (Histamine-releasing factor) (HRF)			
gene; 5'-nucleotidase, ecto			
6qEST; NT5E	3	3	2
gene; transcribed sequence with strong similarity to protein Sorting nexin 14			
6qEST; SNX14	2	3	2
6qEST; Stratagene colon cDNA clone; transcribed sequence with weak similarity to hypothetical protein FLJ20378	3	3	3
6qESTs; SYNCRIP	2	3	3
gene; synaptotagmin binding, cytoplasmic RNA interacting protein			
6qESTs; transcribed sequence with weak similarity to protein KPY2 pyruvate kinase, M2 isozyme	2	3	3
6qEST; Soares_fetal_lung_NbHL19W cDNA clone	5	5	5
pseudogene; similar to 60S ribosomal protein L7			
predicted gene; similar to NADH-ubiquinone oxidoreductase 13 kDa-B subunit (Complex I-13Kd B) (Complex I subunit B13)			
microsatellite marker			
microsatellite marker			
microsatellite marker			
gene; 5-hydroxytryptamine (serotonin) receptor 1E			
gene; glycoprotein hormones (Chorionic gonadotropin), alpha polypeptide			

Entry	6q band	Nucleotide Position	Marker	Uni(Gene/STS) ID	GenBank Accession
E		87790825-87791691	EST 1g3	UniGene: Hs.119689	AA677403
S	6q15	87827007-87830643	LOC377412	na	na
G	6q15	87857091-87967115	ZNF292	UniGene: Hs.448341	XM_048070
E		87875075-87968770	ESTs 1e10, 1c7, 3a7	UniGene: Hs.448341	N33849, R81383, R39115
M		87878779-87878879	D6S1327	UniSTS: 20543	G02953
E		87889985-87890098	EST 1g2	na	AA704826
E		87890552-87890701	ESTs 1d10, 3b5	na	H50661
E		87941729-87941829	EST 1c5	UniGene: Hs.28455	R63899
M		87968649-87968770	D6S1893	UniSTS: 56266	G07043,T60850,Z22792
P	6q15	87989555-88034592	LOC375519	UniGene: Hs.254280	XM_098449
M		88019427-88019633	D6S1297	UniSTS: 64453	G02965
G	6q15-q16.1	88042321-88047070	C6orf162	UniGene: Hs.70769	NM_020425
E		88045592-88047541	ESTs 1g9, 3e1, 2a9	UniGene: Hs.70769	AA120914,AA676890,AA876269
G	6q15	88061458-88070710	C6orf163	UniGene: Hs.446115	XM_116497
E		88111569-88112336	EST 1c6	UniGene: Hs.28327	BX118928, R62370, R72993
G	6q15	88113327-88141371	C6orf165	UniGene: Hs.82921	NM_178823
S	6q14.3-q15	88131109-88131402	TAF13P	na	NG_001573
M		88153155-88153354	D6S1004	UniSTS: 72009	G08497
E		88169513-88169669	EST 3f9	UniGene: Hs.82921	AI340927
G	6q15	88178291-88217643	SLC35A1	UniGene: Hs.82921	BC017807
E		88184482-88184581	EST 1e8	na	H89730
E		88217324-88217444	EST 3g8	UniGene: Hs.511750	AI687708
G	6q16.1	88219699-88295317	RARSL	UniGene: Hs.432918	NM_020320
G	6q14.3-q16.1	88295439-88372762	ORC3L	UniGene: Hs.312419	AF135044
E		88372399-88372569	EST 1h2	UniGene: Hs.312419	AA194876
G	6q15	88380571-88407523	C6orf166	UniGene: Hs.201864	NM_018064
M		88507733-88507861	D6S1595	UniSTS: 17985	Z52894
M		88530265-88530369	D6S496	UniSTS: 11304	G09249
G	6q16.1	88753103-88772146	SPACA1	UniGene: Hs.161241	NM_030960
G	6q14-q15	88845182-88871158	CNR1	UniGene: Hs.75110	NM_001840, NM_016083, NM_033181
E		88845225-88845455	EST 3f2	UniGene: Hs.75110	AI149484
M		88846862-88847114	D6S1909	UniSTS: 24798	G05433
P	6q16.1	88970614-88974176	LOC377413	na	XM_352619
S	6q13	88981075-88982735	ACTBP8	na	NG_000840
E		89175693-89175842	EST 3a1	UniGene: Hs.485921	T57514
M		89208437-89208709	AFMB298ZG9	UniSTS: 44384	Z67582
G	6q16	89315595-89668900	RNGTT	UniGene: Hs.27345	NM_003800
E		89315701-89315833	ESTs 1h9, 3b8	UniGene: Hs.27345	AA425882, H70610
E		89466212-89466437	ESTs 1g11, 3d6	UniGene: Hs.423277	AA156314
M		89556174-89556416	D6S993	UniSTS: 70265	G02948
M		89723479-89723592	D6S1644	UniSTS: 51354	Z53589
G	6q16.1	89786093-89790475	PNRC1	UniGene: Hs.75969	NM_006813
G	6q16.1	89803813-89823330	SRrp35	UniGene: Hs.254414	BC021715
G	6q16.1	89843584-89867578	ACY1L2	UniGene: Hs.356247	BC035036
E		89869996-89870141	ESTs 1d8, 3b2	UniGene: Hs.356247	H43330, AI521245
E		89882871-89884065	ESTs 3c8, 3h3	UniGene: Hs.437745	AA018612, AI701128
G	6q13-q16.3	89883568-89922656	GABRR1	UniGene: Hs.437745	M62400
G	6q13-q16.3	89962835-90020550	GABRR2	UniGene: Hs.99927	M86868
E		90032470-90034557	ESTs 1h3, 1f11, 3g4, 3d1	UniGene: Hs.184325	AA256528, AA074480, AI375320, AA704947
G	6q16.1	90034303-90058028	UBE2J1	UniGene: Hs.184325	NM_016336
E		90063149-90063337	ESTs 1d6, 3f10	UniGene: Hs.31712	H19992, AI347780
E		90068926-90069079	EST 3g3	UniGene: Hs.269028	AI376253
G	6q15-q16	90073190-90117361	RRAGD	UniGene: Hs.238679	AF272036
M		90081531-90081681	AFM185XD10	UniSTS: 13874	M37197, Z67793
E		90114796-90115130	EST 1b2	na	R40888
E		90270548-90337620	ESTs 3e3, 3a9, 3c10	UniGene: Hs.30991	AA620807, R59489, AA502518, AA025623

Description	Dot Blot Expression		
	Per.Ret.	Fovea	RPE
6qEST; CGA	2	2	2
pseudogene; similar to Reticulocalbin 1 precursor			
gene; zinc finger protein 292			
6qESTs; ZNF292	4	4	3
microsatellite marker			
6qEST; Soares_fetal_liver_spleen_1NFLS_S1 cDNA clone	2	2	2
6qESTs; Soares fetal liver spleen 1NFLS cDNA clone	3	3	3
6qEST; Soares placenta Nb2HP cDNA clone	4	4	2
microsatellite marker			
predicted gene; LOC153936: moderately similar to Gap junction beta-5 protein; LOC375519: similar to connexin25			
microsatellite marker			
gene; chromosome 6 open reading frame 162			
6qESTs; C6orf162	2	2	2
gene; chromosome 6 open reading frame 163			
6qEST; Soares placenta Nb2HP cDNA clones	5	5	5
gene; chromosome 6 open reading frame 165; hypothetical protein FLJ25974			
pseudogene; TAF13 RNA polymerase II, TATA box binding protein (TBP)-associated factor 13			
microsatellite marker			
6qEST; C6orf165	5	5	5
gene; solute carrier family 35 (CMP-sialic acid transporter), member A1			
6qEST; Soares fetal liver spleen 1NFLS cDNA clone	4	4	4
6qEST; transcribed sequence with moderate similarity to protein JC5023 CMP-sialic acid transporter	1	3	4
gene; arginyl-tRNA synthetase-like			
gene; origin recognition complex, subunit 3-like (yeast)			
6qEST; ORC3L	2	2	2
gene; chromosome 6 open reading frame 166; hypothetical protein FLJ10342			
microsatellite marker			
microsatellite marker			
gene; sperm acrosome associated 1			
gene; cannabinoid receptor 1 (brain)			
6qEST; CNR1	1	2	3
microsatellite marker			
predicted gene; similar to ribosomal protein S14			
pseudogene; actin, beta pseudogene 8			
6qEST; Stratagene ovary cDNA clone	2	3	3
microsatellite marker			
gene; RNA guanylyltransferase and 5'-phosphatase			
6qESTs; RNGTT	3	3	3
6qESTs; Homo sapiens transcribed sequence with weak similarity to hypothetical protein FLJ20489	5	5	5
microsatellite marker			
microsatellite marker			
gene; proline-rich nuclear receptor coactivator 1			
gene; serine-arginine repressor protein (35 kDa)			
gene; aminoacylase 1-like 2			
6qESTs; ACY1L2	3	3	3
6qESTs; GABRR1	2	3	3
gene; gamma-aminobutyric acid (GABA) receptor, rho 1			
gene; gamma-aminobutyric acid (GABA) receptor, rho 2			
6qESTs; UBE2J1	2	3	3
gene; ubiquitin-conjugating enzyme E2, J1 (UBC6 homolog, yeast)			
6qESTs; Soares adult brain N2b5HB55Y, NCI_CGAP_Kid5 cDNA clones	5	5	5
6qEST; cDNA FLJ45600 fis, clone BRTHA3020369	2	3	2
gene; Ras-related GTP binding D			
microsatellite marker			
6qEST; Soares infant brain 1NIB cDNA clone	1	1	2
6qESTs; ANKRD6	5	5	5

Entry	6q band	Nucleotide Position	Marker	Uni(Gene/STS) ID	GenBank Accession
G	6q14.2-q16.1	90272147-90339151	ANKRD6	UniGene: Hs.30991	AB023174
G	6q14.2-q16.1	90342016-90344062	DJ122O8.2	UniGene: Hs.483841	BC009782
G	6q15	90348829-90525040	MDN1	UniGene: Hs.181334	AF503925
G	6q16.1	90535256-90579752	CASP8AP2	UniGene: Hs.122843	NM_012115
E		90579520-90579661	EST 2a6	UniGene: Hs.122843	AA765960
M		90591268-90591386	D6S1613	UniSTS: 14461	Z53157
E		90706363-90706837	EST 3b6	na	H57955
M		90799754-90799903	D6S1506	UniSTS: 33518	G04930
E		99329745-99330004	EST 2b6	UniGene: Hs.2229	Z11933, AI199744
E		143725691-143725817	EST 3c3	UniGene: Hs.50847	W07262

Description	Dot Blot Expression		
	Per.Ret.	Fovea	RPE
gene; ankyrin repeat domain 6; diversin			
gene; hypothetical protein dJ122O8.2			
gene; midasin homolog (yeast)			
gene; CASP8 associated protein 2			
6qEST; CASP8AP2	3	3	2
microsatellite marker			
6qEST; Soares fetal liver spleen 1NFLS cDNA clone	3	3	3
microsatellite marker			
6qEST; POU3F4: POU domain, class 3, transcription factor 4	3	4	2
6qEST; Bento Soares Infant Brain, fetal_lung_NbHL19W cDNA clones	1	2	2

Figure 3.3 Characterization of various chromosome 6q12-q16 retina/RPE-expressed EST clones by northern blot analysis. Initial dot blots indicating expression in human peripheral retina, fovea, and RPE tissue are shown on the left. Retina/RPE northern blots contain 5 μg of total RNA per lane. Multiple tissue northern blots contain 2 μg of poly(A)$^+$ RNA per lane. Autoradiographs of the northern blots stripped of the EST probes and then re-probed with human β-actin cDNA as a loading control are also shown. Estimated sizes of various transcripts detected by the EST probes are indicated on the right. *Per.Ret.*, peripheral retina; *Sk. Muscle*, skeletal muscle; *kb*, kilobases.

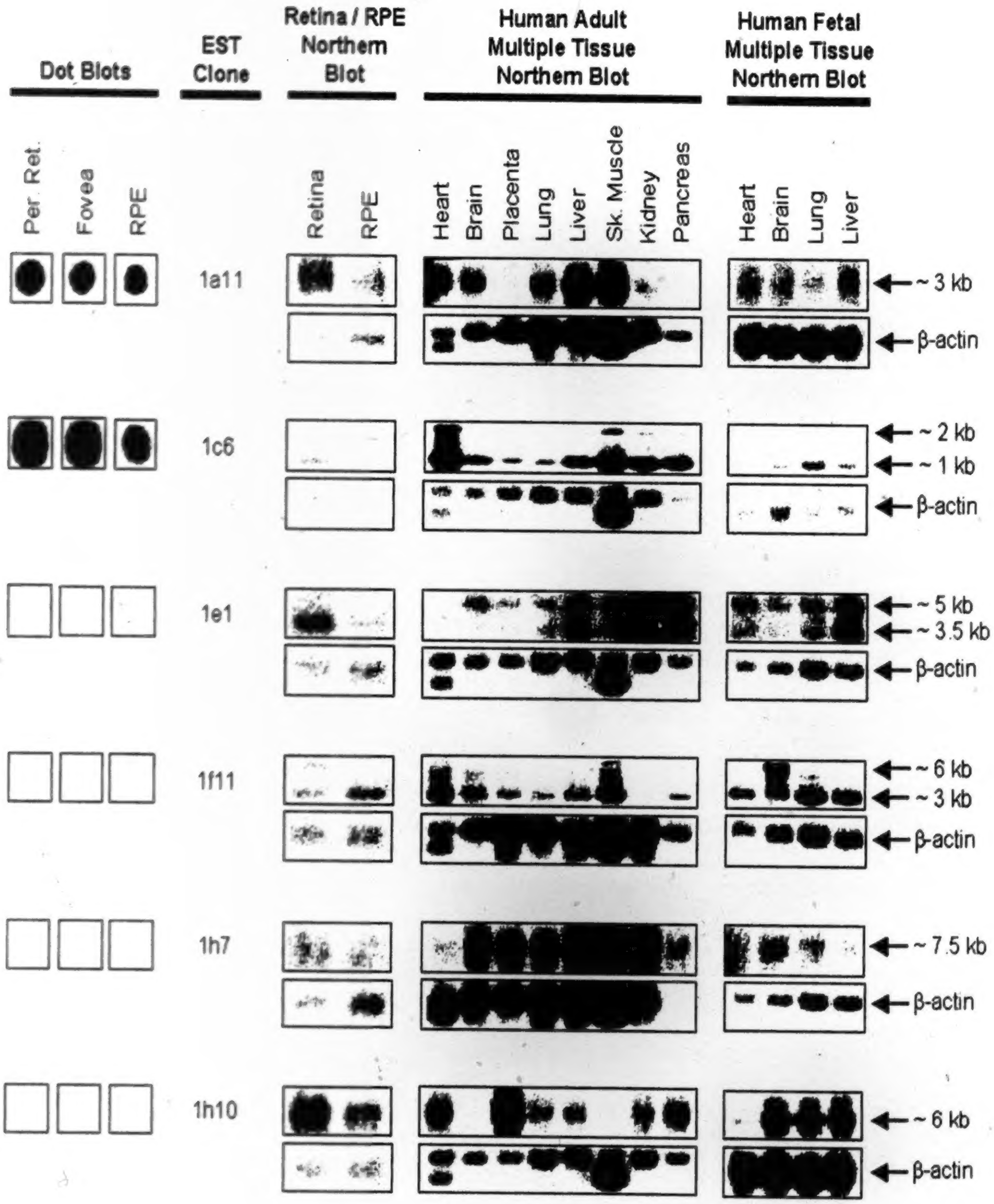
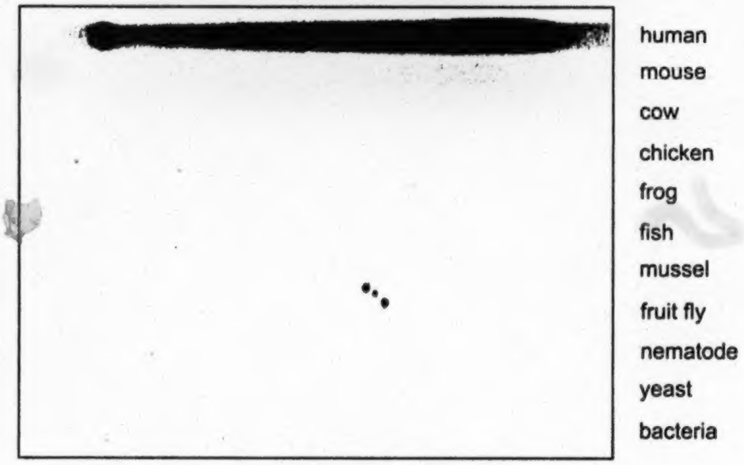
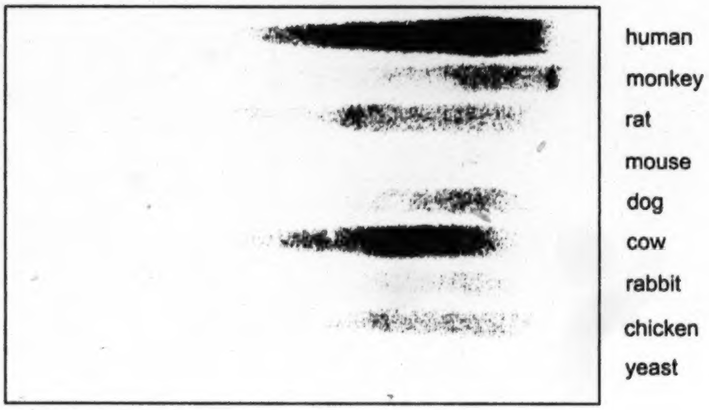

Dot Blots
EST Clone
Retina / RPE Northern Blot
Human Adult Multiple Tissue Northern Blot
Human Fetal Multiple Tissue Northern Blot
Per. Ret.
Fovea
RPE
Retina
RPE
Heart
Brain
Placenta
Lung
Liver
Sk. Muscle
Kidney
Pancreas
Heart
Brain
Lung
Liver
1a11
~ 3 kb
β-actin
1c6
~ 2 kb
~ 1 kb
β-actin
1e1
~ 5 kb
~ 3.5 kb
β-actin
1f11
~ 6 kb
~ 3 kb
β-actin
1h7
~ 7.5 kb
β-actin
1h10
~ 6 kb
β-actin

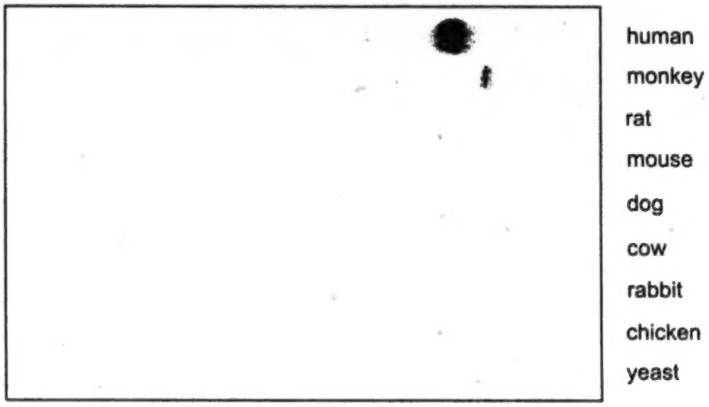
Figure 3.4 Assessment of interspecies sequence conservation of individual retina/RPE-expressed EST clones by Southern blot analysis. Blots contain 8 μg of *Eco*RI-digested genomic DNA extracted from the indicated species. Hybridizations were performed at 42°C with the corresponding EST cDNA clones as probes.


human
mouse
cow
chicken
frog
fish
mussel
fruit fly
nematode
yeast
bacteria

1a11

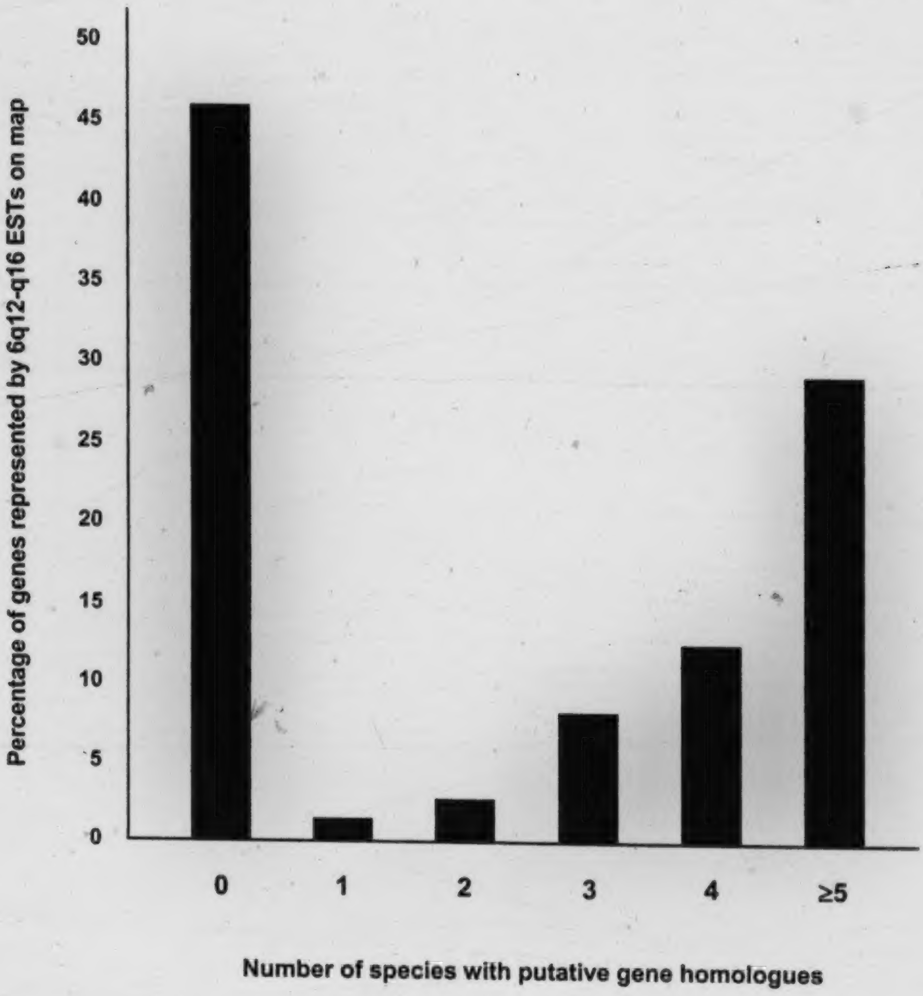

human
monkey
rat
mouse
dog
cow
rabbit
chicken
yeast

1c6


human
monkey
rat
mouse
dog
cow
rabbit
chicken
yeast

1e1

Figure 3.5 Bioinformatic assessment of evolutionary conservation of retinal/RPE-expressed ESTs mapping to chromosome 6q12-q16. Bars on the graph depict the proportion of genes represented by chromosome 6q12-q16 EST clones for which homologues have been identified in the indicated number of non-human species based on *in silico* detection of conserved transcript or annotated genome sequences.


Percentage of genes represented by 6q12-q16 ESTs on map
50
45
40
35
30
25
20
15
10
5
0
0
1
2
3
4
≥5
Number of species with putative gene homologues

Chapter 4

Identification and Characterization of *C6orf37*, a Novel Candidate Human Retinal Disease Gene on Chromosome 6q14

A version of this chapter has been published.
Lagali PS, Kakuk LE, Griesinger IB, Wong PW, Ayyagari R.
Identification and Characterization of *C6orf37*, a Novel Candidate Human Retinal Disease Gene on Chromosome 6q14. *Biochemical and Biophysical Research Communications* 2002;293(1):356-365.

4.1 INTRODUCTION

Human chromosome 6 harbors a wide array of retinal disease loci.[1] At least eight different loci for retinal diseases alone have been localized to the long arm of human chromosome 6 in the region 6q14-q16.[2-11] The genes associated with two of these genetic lesions have recently been identified,[12,13] however those underlying the remaining retinal diseases mapped to this region have yet to be defined. Within the last decade and a half, a number of retinal disease genes have been identified through a straightforward positional cloning approach.[14-19] Rapid advances in the elucidation and annotation of the human genome sequence have enabled a complementary approach to the cloning of disease genes, involving both classical genetic linkage studies combined with sequence database mining for the identification of putative candidate genes in a genomic region of interest. This has led to the discovery in recent years of several genes implicated in hereditary retinopathies.[12,13,20-23] Given the high density of retinal disease loci that have been mapped to human chromosome 6q14-q16, the current goal is to define potential candidate genes in this interval that can be tested for mutations for each genetic disorder. A feasible approach is to identify retina-expressed genes within the 6q14-q16 region and then to assess them for the presence of mutations in affected individuals in families with retinal diseases mapping to this interval. Toward this end, a novel human gene located on chromosome 6q14 that is expressed in the retina has been isolated and characterized. This gene encodes a novel protein that appears to be evolutionarily conserved across a variety of species. The putative gene product is likely a soluble, globular, cytosolic factor that is predicted to have several phosphorylation sites that are shared among orthologues. This gene has been named *chromosome 6 open reading frame 37* (*C6orf37*) and was tested as a candidate gene for retinal disorders mapping to 6q14-q16.

4.2 MATERIALS AND METHODS

4.2.1 Identification and Analysis of ESTs Localized to 6q14-q16

To identify putative genes present in the genomic interval corresponding to the STGD3/adMD critical region,[7,8] we searched for ESTs mapping to the genomic region flanked by microsatellite markers D6S1596 and D6S1644 using human genome resources (http://www.ncbi.nlm.nih.gov/). Fine mapping of the ESTs was established by screening for homologous sequences using the human chromosome 6 Blast server (http://www.sanger.ac.uk/HGP/Chr6/Chr6_blast_server.shtml). Map positions of ESTs were confirmed by PCR amplification of YAC clones comprising a physical contig spanning the 6q14 region (unpublished data). A total of 139 non-overlapping ESTs localized to the region of interest were commercially obtained (Research Genetics, Inc). Clones corresponding to the ESTs were spotted onto nylon membranes for dot blot analysis. Total RNA was isolated from human peripheral retina and human fovea tissue samples using the Trizol reagent (Life Technologies, Inc., Burlington, ON). Five micrograms of total RNA derived from each tissue were used for cDNA synthesis by reverse transcription using the Life Technologies cDNA Synthesis Kit. [α^{32}P]-dCTP-labeled total cDNA probes were generated by random priming using the Multiprime DNA Labeling System (Amersham Biosciences, Baie d'Urfé, QC) followed by purification with G-50 Sephadex spin columns (Eppendorf-5 Prime, Inc., Boulder, CO). Hybridization of the dot blots was carried out in Hybrisol II hybridization solution (Serologicals Corporation, Norcross, GA) at 65°C for 16 hours using 10^6 cpm/ml of radioactively labeled, purified probes. The membranes were then washed twice for 15 minutes in 2X sodium chloride, tri-sodium citrate buffer (SSC) at 65°C, once for 30 minutes in 2X SSC, 0.1% sodium dodecyl sulfate (SDS) at 65°C, and once for 10 minutes in 0.1X SSC, 0.1% SDS at 65°C, followed by exposure to Kodak X-Omat film at -70°C for 24 hours.

4.2.2 Isolation of C6orf37 cDNA

A BLAST search[24] for genomic fragments containing the nucleotide sequence for dot blot hybridization-positive EST clone H65844 identified human chromosome 6 genomic phage P1-derived artificial chromosome (PAC) clone dJ991C6 (GenBank accession number AL078599). Analysis of the dJ991C6 sequence for other ESTs in the region of H65844 yielded 17 ESTs (GenBank accession numbers AW243203, BF025838, AW951915, R74266, AW391612, AW246673, AI022319, AW952387, R78250, AV647023, BG620844, T73543, BE782072, BG622030, BG434626, BG621371, AI990503) that formed a minimal overlapping contig of 5591 base pairs (bp), defining the presumed entire *C6orf37* mRNA transcript-encoding region. An open reading frame (ORF) of 1329 base pairs was found within the predicted *C6orf37* mRNA transcript sequence using the ORF Finder program (http://www.ncbi.nlm.nih.gov/gorf/). Five ESTs that encompassed the putative protein-encoding region (AA551498, R69202, R26344, H65844, AA564797) were purchased (Research Genetics, Inc.) and subjected to automated sequencing using an ABI Prism 377 Sequencer (Perkin Elmer). Sequence assembly was carried out using the Sequencher program (Gene Codes Corporation, Ann Arbor, MI), and ORF detection was performed with the DNASTAR software package. The 5'- and 3'-ends of the *C6orf37* cDNA were defined by identifying transcription initiation[25,26] and transcript cleavage motifs[27,28] respectively in the PAC sequences flanking the EST contig sequence (see Results).

4.2.3 C6orf37 DNA and Protein Sequence Analysis

Alignment of the *C6orf37* cDNA sequence with the genomic sequence contained in the PAC clone dJ991C6 revealed the location of intron-exon boundaries. PCR primers flanking the proposed exons were designed and used to amplify and sequence the corresponding sequences from human genomic DNA to confirm the predicted *C6orf37* gene structure. The following primers were generated: Exon1F, 5'-AGTCGTCCCGCACTAACTGC-3'; Exon1R, 5'-AAGCGCAAGCGAGAGTCCCG-3'; Exon2F, 5'-CGCTCCTTTCCCTTTCTTTT-3'; Exon2R, 5'-TTAACCGATGCCCCTCTG-3'; Exon3F, 5'-TTTCTAGGTTTGCTCATTTTGTTTT-3'; Exon3R, 5'-GCTTCCCCACAGGACATTT-3'.

DNA and protein homology searches for *C6orf37* sequences were performed using the BLASTN and BLASTP search programs at the NCBI, respectively.[24] Additional protein sequence characterization was carried out using bioinformatics tools available through the ExPASy Molecular Biology server (http://ca.expasy.org/tools/) and the Canadian Bioinformatics Resource (http://cbr-rbc.nrc-cnrc.gc.ca/).

4.2.4 Zooblot Analysis

A commercially available zooblot containing *Eco*RI-digested genomic DNA from various species (human, monkey, rat, mouse, dog, cow, rabbit, chicken, yeast) was obtained (BD Biosciences Clontech, Palo Alto, CA) and analyzed by Southern blotting. Hybridizations and washes were performed at 42°C using the above procedure and a radioactively labeled probe corresponding to the *C6orf37* cDNA sequence. The washed blot was exposed to X-ray film at -70°C overnight.

4.2.5 Expression Analysis by Northern Blotting

Northern blots were prepared using five-microgram samples of total RNA isolated from human ocular tissues (retina, retinal pigment epithelium, and choroid). Human fetal and adult multi-tissue northern blots containing two micrograms of poly(A)$^+$ RNA derived from various tissues (heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon, peripheral blood leukocyte) were commercially obtained (BD Biosciences Clontech). Hybridizations and washes of all blots were performed at 65°C using radioactively labeled *C6orf37* and human β-actin cDNA probes. Washed blots were exposed to X-ray film at -70°C for 1-4 days.

4.2.6 In vitro Translation

EST clone AA564797 was used as the template for *in vitro* translation using the TNT Coupled Reticulocyte Lysate System (Promega Corp., Madison, WI) and ^{35}S-methionine. Reaction products were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by enhanced autoradiography. *In vitro* translation of C6orf37 cDNA and subsequent SDS-PAGE was performed by Ms. Laura E. Kakuk (University of Michigan, Ann Arbor, MI).

4.2.7 Mutational Analysis of C6orf37 in STGD3/adMD Pedigrees

PCR reactions to amplify the C6orf37 protein-encoding exons were performed using the Exon2F, Exon2R, Exon3F, and Exon3R primers described above and 100 nanograms of genomic DNA extracted from blood samples obtained from STGD3/adMD family members[7,8] as template in 20-microlitre reactions. PCR cycling parameters were as follows: 1 cycle of denaturation at 94°C for 3 minutes; 35 cycles of denaturation at 94°C for 30 seconds, primer annealing at 55°C for 30 seconds, primer extension at 72°C for 40 seconds; 1 cycle of extension at 72°C for 3 minutes. The amplification products were sequenced using the PCR primers as well as two internal primers for exons 2 and 3 (Exon2INT, 5'-TGAGGTCCAGGTCCTTGTAG-3'; Exon3INT, 5'-TGAAGGGTCATGAGATACTC-3') to detect the presence of sequence alterations in affected individuals from STGD3/adMD families. Sequencing reactions were prepared using the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit (USB Corp., Cleveland, OH). Reaction products were electrophoresed on 6% polyacrylamide gels followed by autoradiography at room temperature.

4.3 RESULTS

4.3.1 Isolation of Human C6orf37 cDNA

To identify genes within the STGD3/adMD critical region on chromosome 6q14-q16 that may represent putative candidate genes for the retinal diseases that map to this genomic interval,[2-11] 139 non-overlapping ESTs that have also been reported to be localized to this region were screened for retinal expression. Dot blot analysis of EST H65844 with human peripheral retina and human fovea total cDNA probes indicated expression in both retinal tissues (data not shown). Fine mapping of H65844 on a YAC contig of the 6q14-q16 region confirmed its localization between microsatellite markers D6S463 and D6S1455 (Figure 4.1A - 4.1C). Bioinformatic analysis of the H65844 EST sequence resulted in the identification of numerous ESTs that spanned a 5591-bp segment contained within the chromosome 6q genomic PAC clone dJ991C6. The sequence of 17 of these ESTs (AW243203, BF025838, AW951915, R74266, AW391612, AW246673, AI022319, AW952387, R78250, AV647023, BG620844, T73543, BE782072, BG622030, BG434626, BG621371, AI990503) could be assembled to form a minimal contig comprising the entire 5591-bp expressed sequence. Analysis of the transcript sequence revealed the presence of a 1329-bp open reading frame. DNA sequencing and sequence assembly of five ESTs (AA551498, R69202, R26344, H65844, AA564797) that encompassed the predicted protein-encoding region of the newly identified gene, which we named *C6orf37*, was subsequently performed. Minor discrepancies were detected between the contig assembled by sequencing the five EST clones and the mRNA sequence predicted from the reported PAC clone sequence, including the absence of a 15-nucleotide (and therefore 5-amino acid) repeat, and four occurrences of G↔C switched nucleotide assignments resulting in two amino acid changes. However the reading frame was not altered and no significant changes to the overall protein sequence were noted. Additional changes found correspond to exon 1 sequence alterations that do not affect the translation product, such as the absence of three individual nucleotides and a T↔C reversal within the 5'-untranslated region (UTR). The sequence discrepancies detected may represent polymorphisms between different tissue sources or may simply reflect the existence of multiple sequence variants. The results thus indicate a 1314-bp full-length ORF contained within the deduced 5571-bp cDNA sequence for *C6orf37* (Figure 4.2).

In addition to the position of constituent EST sequence boundaries, the 5'-end of the complete cDNA was defined by identifying a canonical TATA box promoter element[25] at nucleotide positions –21 to –15 and a eukaryotic transcription initiator (Inr) sequence[26] in the genomic PAC clone just upstream of the putative *C6or37* transcription start site (+1 position). The 3'-end of the cDNA was delineated by the presence of a poly(A)$^+$ signal[27] at position 5546-5551 just upstream of a consensus mRNA cleavage site[28] in the genomic sequence. Analysis of the complete cDNA sequence with gene prediction software located the ATG start codon of the 1314-bp continuous open reading frame at position 295 and the TAA stop codon at position 1606. The sequence surrounding the proposed initiation codon represents a suitable context for translation initiation in vertebrate mRNAs.[29] An in-frame TGA stop codon was found at position 124, 171 nucleotides upstream of the predicted start codon, suggesting that the 1314-bp sequence does in fact represent a complete ORF.

4.3.2 Determination of C6orf37 Genomic Structure

Alignment of the *C6orf37* expressed sequence with the genomic PAC clone dJ991C6 sequence reveals that the gene is organized into 3 exons. Exon 1 comprises the first 257 nucleotides of the 5'-UTR. Exon 2 contains 574 bp, consisting of the terminal 37 bp of the 5'-UTR followed by the initial 537 bp of the protein-encoding region of the gene. Exon 3 contains 4740 bp, 777 bp of which completes the coding region and the remaining 3963 bp corresponds to the 3'-UTR (Figure 4.1D). The exons are separated by two intervening intron sequences that are 252 bp and 1118 bp in length respectively. All splice junctions conform to consensus sequences for splice donor and acceptor sites.[30] The genomic organization of the *C6orf37* gene was confirmed by using primers that flank the proposed exons to amplify the putative gene regions from human genomic DNA. The resulting amplicons were sequenced and found to match the deduced *C6orf37* gene sequence.

4.3.3 Sequence Conservation of C6orf37

Searches of non-redundant DNA and protein databases with the BLAST algorithms did not identify any significant homologies between *C6orf37* and previously characterized human genes or proteins. Similarities of the proposed *C6orf37* translation product to putative proteins in mouse (72%), *D. melanogaster* (60%), and *C. elegans* (40%) were identified (Figure 4.3), however the function of these proteins is as of yet unknown. Sequence assembly and translation of six overlapping *R. norvegicus* ESTs (AA998465, AA926118, AA933175, AW502512, BF558308, BI294811) displaying significant similarity to the *C6orf37* cDNA sequence also revealed the existence of a rat homologue that shares 91% identity with the human protein (Figure 4.3). In addition, zooblot analysis showed hybridization of *C6orf37* cDNA probes to genomic DNA extracted from a number of other species including chicken, rabbit, cow, and the yeast *Saccharomyces cerevisiae* (data not shown), further suggesting the evolutionary conservation of this gene in both mammalian and non-mammalian species.

4.3.4 C6orf37 Protein and mRNA Sequence Analysis

The 1314-bp ORF encodes a protein of 437 amino acids with an apparent molecular weight of 49.2 kilodaltons (kDa) and an isoelectric point of 5.01. *In vitro* translation using the EST clone AA564797, which was found to encompass the entire *C6orf37* coding sequence, generated a protein product of approximately 49 kDa as predicted (Figure 4.4). Analysis of the C6orf37 amino acid sequence with the NetPhos 2.0 program[31] predicts the presence of several potential phosphorylation sites that are found to be conserved among the different orthologues (Figure 4.3). Protein structure prediction using the PredictProtein software[32] reveals the absence of any putative hydrophobic transmembrane segments and predicts a globular structure, suggesting that the C6orf37 protein is unlikely to be membrane bound. Subcellular localization of the protein predicted by the PSORTII program[33] indicates that it is likely to be found in the cytoplasm due to the lack of any known intracellular organelle targeting signals. Searches of the C6orf37 amino acid sequence against a variety of protein motif databases using the ProfileScan, PROSITE,[34] PRINTScan,[35] BLOCKS,[36] pFAM,[37] and ProDom[38] programs failed to find significant homology to any known protein domains.

Additional examination of the *C6orf37* cDNA sequence reveals the presence of several ATTTA motifs in the lengthy 3'-UTR (Figure 4.2). These AU-rich elements (AREs) are believed to play a role in determining RNA stability in mammalian cells,[39] suggesting that the *C6orf37* transcript may undergo rapid and selective decay.

4.3.5 Tissue Distribution of C6orf37 mRNA

To determine the expression pattern of *C6orf37*, northern blot analysis of human fetal and adult tissues was performed using the EST H65844 sequence as a probe (Figure 4.5). Two transcript sizes were detected, the major one consistent with the proposed full-length 5571-bp *C6orf37* cDNA species, and a second, less abundant and slightly larger transcript possibly arising from alternative splicing. The predominant, approximately 5.6-kb message is expressed in most tissues tested. Strongest expression is found in the ovary, while it is weakly detectable in adult skeletal muscle, thymus, liver, lung, and brain. Interestingly, the *C6orf37* transcripts are prominently expressed in fetal liver, lung, and brain. In addition the *C6orf37* 5.6-kb transcript was observed in adult heart but not in fetal heart tissue. These findings suggest that the *C6orf37* gene may be developmentally regulated in some tissues. Within the immediate region of the retina, the major *C6orf37* transcript was detected in the neural retina but not in the adjacent retinal pigment epithelium, nor in the choroid (Figure 4.5). The initial dot blot analysis performed for two non-overlapping ESTs corresponding to the *C6orf37* cDNA sequence (H65844 and H92739) indicates that the gene is expressed in both the foveal and peripheral regions of the adult human retina (data not shown).

4.3.6 Exclusion of C6orf37 as the STGD3/adMD Causative Gene

To test whether *C6orf37* represents the STGD3/adMD disease gene, the protein-encoding exons and intron-exon boundaries were screened for the presence of mutations among affected individuals of families diagnosed with Stargardt-like macular dystrophy/autosomal dominant atrophic macular degeneration in which the putative disease gene(s) have been mapped to the 6q14-q16 region.[7,8] No sequence alterations were found in affected versus unaffected individuals (data not shown), and thus *C6orf37* was excluded as the gene that underlies the condition seen in these families.

4.4 DISCUSSION

Through a procedure of genomic and EST sequence database mining coupled with experimental analysis for tissue-selective expression status, this study describes a strategy for the identification of candidate genes for inherited diseases mapping to a genomic interval of interest. By this method, a novel gene was identified that maps to the human chromosome 6q14-q16 region, a genomic interval harboring multiple retinal disease loci. Isolation and characterization of this gene, named *C6orf37*, is reported here.

The *C6orf37* gene is composed of three exons and two introns, spanning approximately 7 kilobases of genomic DNA. A transcript length of 5571 nucleotides is predicted for the functional *C6orf37* message, encoding a putative protein of 437 amino acids. The gene product appears to share significant homology with predicted proteins in *R. norvegicus*, *M. musculus*, *C. elegans*, and *D. melanogaster*, however the biological role of *C6orf37* or any of its orthologues is at present unknown. No homology of *C6orf37* with any protein of known function was detected, and bioinformatic analysis of the amino acid sequence failed to reveal the presence of any previously characterized functional motifs. However, the extent of sequence conservation observed across a wide range of species suggests that *C6orf37* and its homologous counterparts may represent the first-identified members of a novel protein family with an important cellular role.

Clues to the nature of this role were gained from additional sequence analysis of the *C6orf37* gene. A number of conserved potential phosphorylation sites were detected in the primary structure of the C6orf37 protein, which may be involved in post-translational activation or regulation of its function. The lack of any transmembrane segments or subcellular organelle targeting signals and the prediction of its cytoplasmic localization is consistent with its possible role as a soluble intracellular signaling molecule that may serve as a target for the action of kinase and/or phosphatase enzymes. Control of *C6orf37* function may also be imparted at the level of mRNA turnover, as is suggested by the presence of multiple ATTTA sequence motifs in the 3'-

UTR that are known to act as RNA destabilizing elements.[39] Multiple *C6orf37* regulatory mechanisms including potential protein phosphorylation or dephosphorylation and selective transcript degradation may indicate the importance of rapid signal transduction mediated through the *C6orf37* gene product in a particular functional pathway.

Northern blotting of RNA from human ocular tissues confirmed the retinal expression initially detected for *C6orf37* constituent ESTs by dot blot analysis, and further demonstrated preferential expression of the gene within the neural retina over other eye tissues such as the retinal pigment epithelium and the choroid. However, retina-specific expression is not a requisite feature of retinal disease genes, as has been observed for a number of ubiquitously expressed genes that underlie various forms of retinal degeneration,[16-19,40-42] and therefore should not be considered a strict requirement of potential candidates. Additional analysis of human non-ocular tissues suggests that the *C6orf37* gene is also expressed in a wide range of tissue types, consistent with the broad spectrum of tissue sources noted for derivative ESTs. Furthermore, a developmental role is suggested by the differential expression observed between the adult and fetal tissues examined. This may indicate that there are different functions of the C6orf37 protein in different tissues at various stages of development. Studies are underway to extend this finding to human ocular tissues as well as various tissues from non-human species. The northern blots also reveal the presence of at least two *C6orf37* transcripts in several tissue types. The larger, less abundant transcript may be the product of a rare splicing event, and raises the possibility that alternative splicing may be another means by which *C6orf37* gene expression is regulated.

Satisfying the criteria of mapping to the chromosome 6q14-q16 interval and clearly displaying retinal expression, the *C6orf37* candidate gene was screened for mutations in affected individuals of STGD3/adMD families that have previously been described and for which the disease loci were mapped to 6q14-q16.[7,8] No mutations were detected within the protein-encoding exons of *C6orf37*, leading to its exclusion as the gene associated with the disease in these families. As it turns out, a 5-bp deletion in another gene that was subsequently localized to the same genomic region, *ELOVL4*, was found to define the genetic lesion in these families.[12] Given the mapping and expression status of *C6orf37* and *ELOVL4*, they should both be tested as putative candidates for the other retinal dystrophies that are known to localize within the 6q14-q16 region. Identification and characterization of additional retina-expressed genes located in this chromosomal interval will be useful in elucidating the myriad of possible mutations that give rise to the retinal disorders that map to this unique region of the human genome.

4.5 ACKNOWLEDGEMENTS

STGD3/adMD patients were provided by Dr. Paul A. Sieving and Dr. Ian M. MacDonald. The patients and their families are thanked for their participation in these studies. This work was supported by The Foundation Fighting Blindness USA, NEI Vision Core Grant EY0-7003, RPB Harrington Scholar Award (RA); Foundation Fighting Blindness-Canada (PWW); E.A.Baker Foundation of the Canadian National Institute for the Blind (PWW, PSL); Canadian Institutes of Health Research Doctoral Research Award (PSL).

4.6 REFERENCES

1. Daiger SP. RetNet, the Retinal Information Network http://www.sph.uth.tmc.edu/Retnet/ .

2. Small KW, Weber JL, Roses A, Lennon F, Vance JM, Pericak-Vance MA. (1992) North Carolina macular dystrophy is assigned to chromosome 6. *Genomics* 13:681-685.

3. Stone EM, Nichols BE, Kimura AE, Weingeist TA, Drack A, Sheffield VC. (1994) Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. *Arch. Ophthalmol.* 112:765-772.

4. Kelsell RE, Godley BF, Evans K, Tiffin PA, Gregory CY, Plant C, Moore AT, Bird AC, Hunt DM. (1995) Localization of the gene for progressive bifocal chorioretinal atrophy (PBCRA) to chromosome 6q. *Hum. Mol. Genet.* 4:1653-1656.

5. Kelsell RE, Gregory-Evans K, Gregory-Evans CY, Holder GE, Jay MR, Weber BH, Moore AT, Bird AC, Hunt DM. (1998) Localization of a gene (CORD7) for a dominant cone-rod dystrophy to chromosome 6q. *Am. J. Hum. Genet.* 63:274-279.

6. Ruiz A, Borrego S, Marcos I, Antinolo G. (1998) A major locus for autosomal recessive retinitis pigmentosa on 6q, determined by homozygosity mapping of chromosomal regions that contain gamma-aminobutyric acid-receptor clusters. *Am. J. Hum. Genet.* 62:1452-1459.

7. Griesinger IB, Sieving PA, Ayyagari R. (2000) Autosomal dominant macular atrophy at 6q14 excludes CORD7 and MCDR1/PBCRA loci. *Invest. Ophthalmol. Vis. Sci.* 41:248-255.

8. Lagali PS, MacDonald IM, Griesinger IB, Chambers ML, Ayyagari R, Wong PW. (2000) Autosomal dominant Stargardt-like macular dystrophy segregating in a large Canadian family. *Can. J. Ophthalmol.* 35:315-324.

9. Dharmaraj S, Li Y, Robitaille JM, Silva E, Zhu D, Mitchell TN, Maltby LP, Baffoe-Bonnie AB, Maumenee IH. (2000) A novel locus for Leber congenital amaurosis maps to chromosome 6q. *Am. J. Hum. Genet.* 66:319-326.

10. Kniazeva M, Traboulsi EI, Yu Z, Stefko ST, Gorin MB, Shugart YY, O'Connell JR, Blaschak CJ, Cutting G, Han M, Zhang K. (2000) A new locus for dominant drusen and macular degeneration maps to chromosome 6q14. *Am. J. Ophthalmol.* 130:197-202.

11. van Lith-Verhoeven JJ, Hoyng CB, van den Helm B, Deutman AF, Brink HM, Kemperman MH, de Jong WH, Kremer H, Cremers FP. (2004) The benign concentric annular macular dystrophy locus maps to 6p12.3-q16. *Invest. Ophthalmol. Vis. Sci.* 45:30-5.

12. Zhang K, Kniazeva M, Han M, Li W, Yu Z, Yang Z, Li Y, Metzker ML, Allikmets R, Zack DJ, Kakuk LE, Lagali PS, Wong PW, MacDonald IM, Sieving PA, Figueroa DJ, Austin CP, Gould RJ, Ayyagari R, Petrukhin K. (2001) A 5-bp deletion in *ELOVL4* is associated with two related forms of autosomal dominant macular dystrophy. *Nat. Genet.* 27:89-93.

13. Johnson S, Halford S, Morris AG, Patel RJ, Wilkie SE, Hardcastle AJ, Moore AT, Zhang K, Hunt DM. (2003) Genomic organisation and alternative splicing of human *RIM1*, a gene implicated in autosomal dominant cone-rod dystrophy (CORD7). *Genomics* 81:304-314.

14. Lee WH, Bookstein R, Hong F, Young LJ, Shew JY, Lee EY. (1987) Human retinoblastoma susceptibility gene: cloning, identification, and sequence. *Science* 235:1394-1399.

15. Cremers FP, van de Pol DJ, van Kerkhoff LP, Wieringa B, Ropers HH. (1990) Cloning of a gene that is rearranged in patients with choroideraemia. *Nature* 347:674-677.

16. Meindl A, Dry K, Herrmann K, Manson F, Ciccodicola A, Edgar A, Carvalho MR, Achatz H, Hellebrand H, Lennon A, Migliaccio C, Porter K, Zrenner E, Bird A, Jay M, Lorenz B, Wittwer B, D'Urso M, Meitinger T, Wright A. (1996) A gene (*RPGR*) with homology to the RCC1 guanine nucleotide exchange factor is mutated in X-linked retinitis pigmentosa (RP3). *Nat. Genet.* 13:35-42.

17. Roepman R, van Duijnhoven G, Rosenberg T, Pinckers AJ, Bleeker-Wagemakers LM, Bergen AA, Post J, Beck A, Reinhardt R, Ropers HH, Cremers FP, Berger W.

(1996) Positional cloning of the gene for X-linked retinitis pigmentosa 3: homology with the guanine-nucleotide-exchange factor RCC1. *Hum. Mol. Genet.* 5:1035-1041.

18. Schwahn U, Lenzner S, Dong J, Feil S, Hinzmann B, van Duijnhoven G, Kirschner R, Hemberger M, Bergen AA, Rosenberg T, Pinckers AJ, Fundele R, Rosenthal A, Cremers FP, Ropers HH, Berger W. (1998) Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nat. Genet.* 19:327-332.

19. Weber BH, Vogt G, Pruett RC, Stohr H, Felbor U. (1994) Mutations in the tissue inhibitor of metalloproteinases-3 (TIMP3) in patients with Sorsby's fundus dystrophy. *Nat. Genet.* 8:352-356.

20. Sauer CG, Gehrig A, Warneke-Wittstock R, Marquardt A, Ewing CC, Gibson A, Lorenz B, Jurklies B, Weber BH. (1997) Positional cloning of the gene associated with X-linked juvenile retinoschisis. *Nat. Genet.* 17:164-170.

21. Petrukhin K, Koisti MJ, Bakall B, Li W, Xie G, Marknell T, Sandgren O, Forsman K, Holmgren G, Andreasson S, Vujic M, Bergen AA, McGarty-Dugan V, Figueroa D, Austin CP, Metzker ML, Caskey CT, Wadelius C. (1998) Identification of the gene responsible for Best macular dystrophy. *Nat. Genet.* 19:241-247.

22. Sohocki MM, Bowne SJ, Sullivan LS, Blackshaw S, Cepko CL, Payne AM, Bhattacharya SS, Khaliq S, Qasim Mehdi S, Birch DG, Harrison WR, Elder FF, Heckenlively JR, Daiger SP. (2000) Mutations in a new photoreceptor-pineal gene on 17p cause Leber congenital amaurosis. *Nat.Genet.* 24:79-83.

23. Nishimura DY, Searby CC, Carmi R, Elbedour K, Van Maldergem L, Fulton AB, Lam BL, Powell BR, Swiderski RE, Bugge KE, Haider NB, Kwitek-Black AE, Ying L, Duhl DM, Gorman SW, Heon E, Iannaccone A, Bonneau D, Biesecker LG, Jacobson SG, Stone EM, Sheffield VC. (2001) Positional cloning of a novel gene on chromosome 16q causing Bardet-Biedl syndrome (BBS2) *Hum. Mol. Genet.* 10:865-874.

24. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.

25. McKnight SL, Kingsbury R. (1982) Transcriptional control signals of a eukaryotic protein-coding gene. *Science* 217:316-324.

26. Javahery R, Khachi A, Lo K, Zenzie-Gregory B, Smale ST. (1994) DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol. Cell. Biol.* 14:116-127.

27. Proudfoot N. (1991) Poly(A) signals. *Cell* 64:671-674.

28. Chen F, MacDonald CC, Wilusz J. (1995) Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res.* 23:2614-2620.

29. Kozak M. (1996) Interpreting cDNA sequences: some insights from studies on translation. *Mamm. Genome* 7:563-574.

30. Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA. (1986) Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* 55:1119-1150.

31. Blom N, Gammeltoft S, Brunak S. (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* 294:1351-1362.

32. Rost B. (1996) PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol.* 266:525-539.

33. Nakai K, Kanehisa M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14:897-911.

34. Hofmann K, Bucher P, Falquet L, Bairoch A. (1999) The PROSITE database, its status in 1999. *Nucleic Acids Res.* 27:215-219.

35. Attwood TK, Croning MD, Flower DR, Lewis AP, Mabey JE, Scordis P, Selley JN, Wright W. (2000) PRINTS-S: the database formerly known as PRINTS. *Nucleic Acids Res.* 28:225-227.

36. Henikoff JG, Greene EA, Pietrokovski S, Henikoff S. (2000) Increased coverage of protein families with the blocks database servers. *Nucleic Acids Res.* 28:228-230.

37. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. (2000) The Pfam protein families database. *Nucleic Acids Res.* 28:263-266.

38. Corpet F, Servant F, Gouzy J, Kahn D. (2000) ProDom and ProDom-CG: tools for protein domain analysis and whole genome comparisons. *Nucleic Acids Res.* 28:267-269.

39. Xu N, Chen CY, Shyu AB. (1997) Modulation of the fate of cytoplasmic mRNA by AU-rich elements: key sequence features controlling mRNA deadenylation and decay. *Mol. Cell. Biol.* 17:4611-4621.

40. Stone EM, Lotery AJ, Munier FL, Heon E, Piguet B, Guymer RH, Vandenburgh K, Cousin P, Nishimura D, Swiderski RE, Silvestri G, Mackey DA, Hageman GS, Bird AC, Sheffield VC, Schorderet DF. (1999) A single *EFEMP1* mutation associated with both Malattia Leventinese and Doyne honeycomb retinal dystrophy. *Nat. Genet.* 22:199-202.

41. Gal A, Li Y, Thompson DA, Weir J, Orth U, Jacobson SG, Apfelstedt-Sylla E, Vollrath D. (2000) Mutations in *MERTK*, the human orthologue of the RCS rat retinal dystrophy gene, cause retinitis pigmentosa. *Nat. Genet.* 26:270-271.

42. McKie AB, McHale JC, Keen TJ, Tarttelin EE, Goliath R, van Lith-Verhoeven JJ, Greenberg J, Ramesar RS, Hoyng CB, Cremers FP, Mackey DA, Bhattacharya SS, Bird AC, Markham AF, Inglehearn CF. (2001) Mutations in the pre-mRNA splicing factor gene *PRPC8* in autosomal dominant retinitis pigmentosa (RP13). *Hum. Mol. Genet.* 10:1555-1562.

Figure 4.1 Cloning strategy used to identify and define the human *C6orf37* gene. Refer to text for details. *A*, Representation of the STGD3/adMD interval relative to human chromosome 6q microsatellite marker positions. *cen*, centromere; *tel*, telomere. *B*, Genomic PAC clone dJ991C6 (GenBank accession number AL078599) mapping to the STGD3/adMD interval. Xs denote the presence of the corresponding chromosome 6q microsatellite markers shown in Panel *A*. *C*, Human ESTs spanning the *C6orf37* open reading frame. Stippled black rectangles represent EST-spanning regions, while intervening horizontal lines indicate the locations of gaps in the alignment of the EST sequences with the genomic PAC sequence. GenBank accession numbers for each of the EST clones are given. *D*, Genomic organization of the *C6orf37* gene. Rectangles represent exons and intervening horizontal lines represent introns. Open rectangles correspond to the open reading frame, while shading denotes untranslated regions (UTR). The lengths of exon sequences are given within the rectangles, and approximate sizes of the introns are indicated below the intervening horizontal lines. The positions of the start (ATG) and stop (TAA) codons of the ORF are shown. *E*, *C6orf37* exon-intron boundary sequences. Nucleotides comprising intron-flanking exon sequences are given in upper case, while terminal portions of the intron sequences are given in lower case.

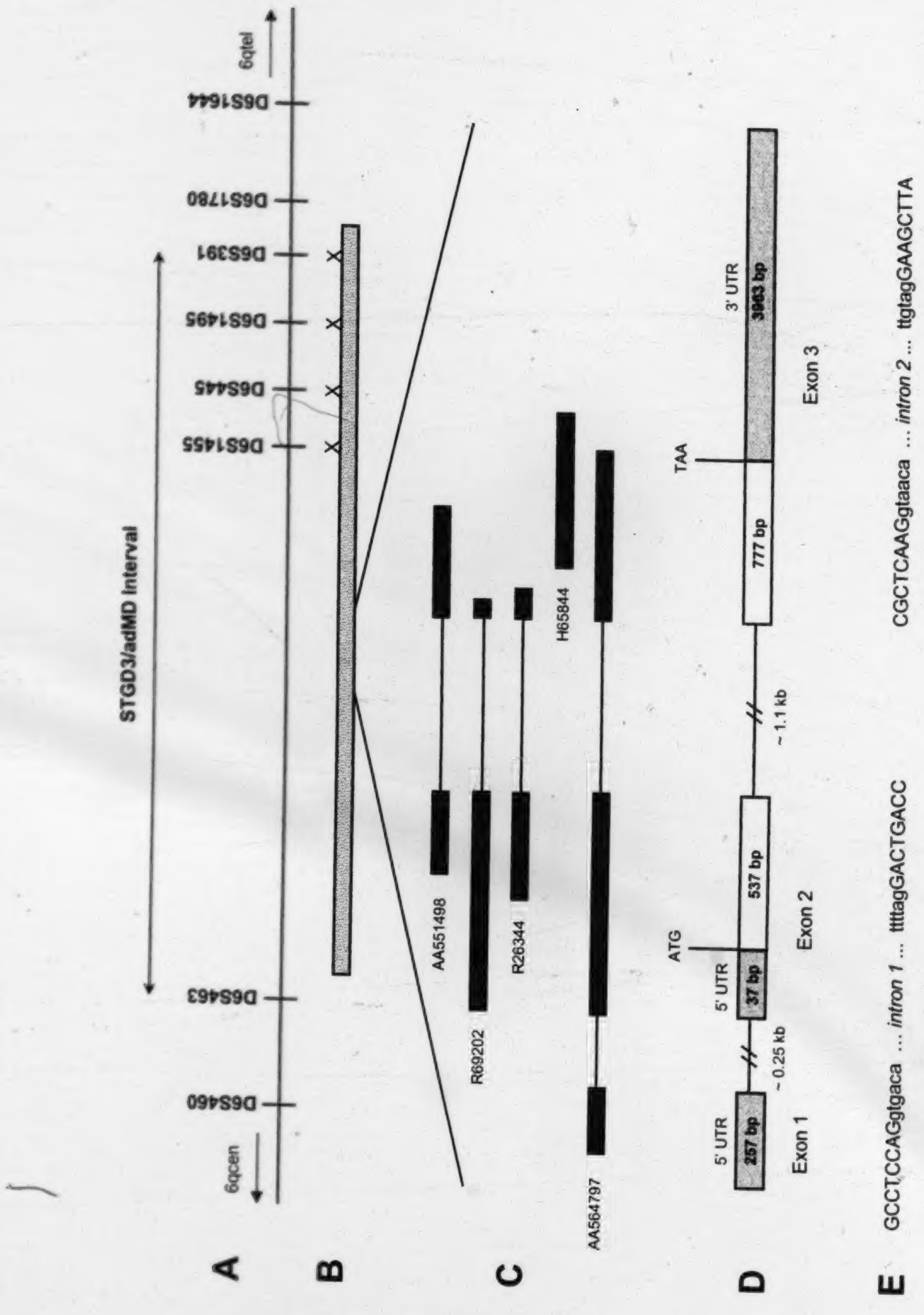
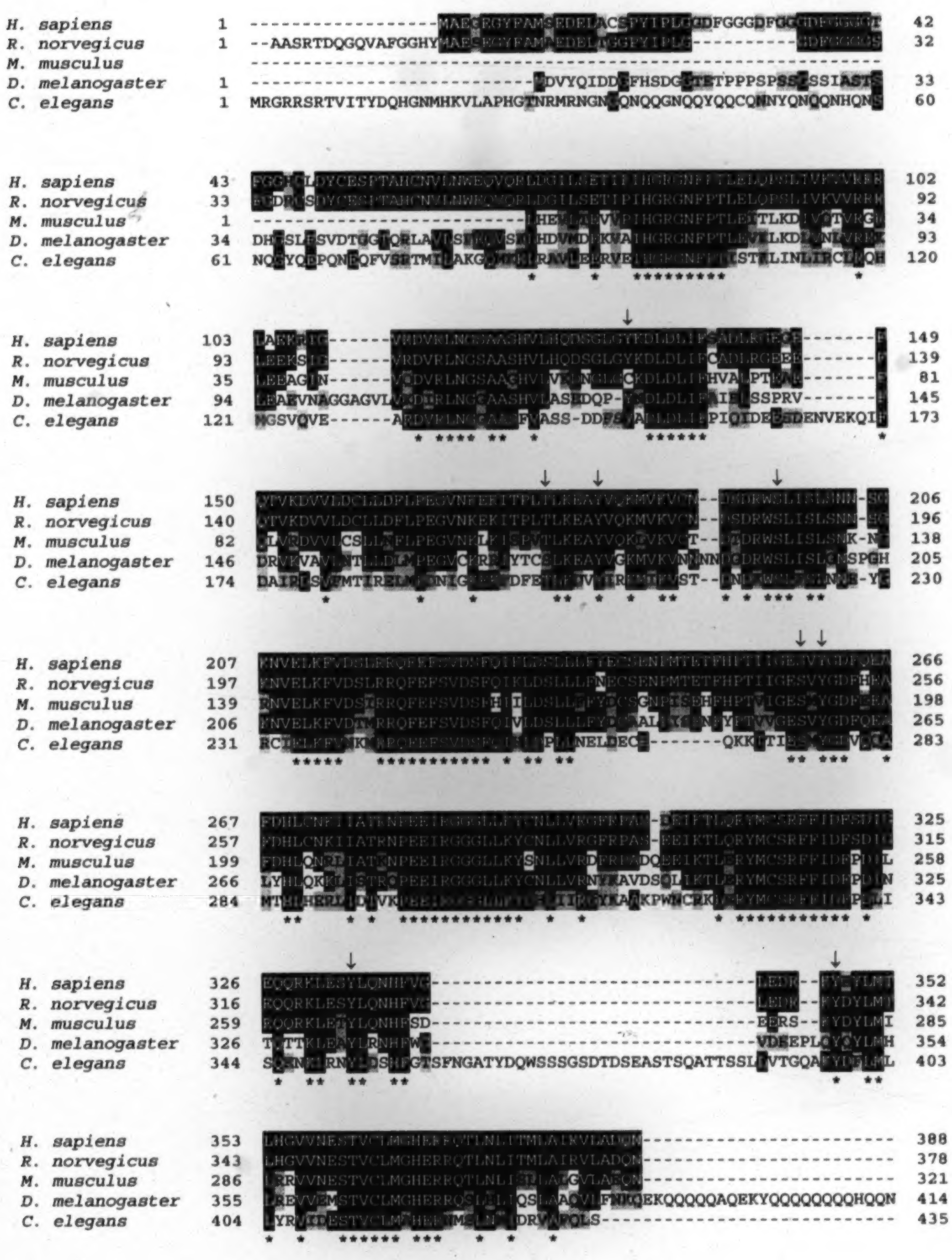

A
B
C
D
E
6qcen
6qtel
D6S460
D6S463
D6S1455
D6S445
D6S1495
D6S391
D6S1780
D6S1644
STGD3/adMD Interval
AA551498
R69202
R26344
H65844
AA564797
5' UTR
257 bp
~ 0.25 kb
Exon 1
5' UTR
37 bp
ATG
537 bp
Exon 2
~ 1.1 kb
777 bp
TAA
3' UTR
3963 bp
Exon 3
GCCTCCAGgtgaca ... intron 1 ... ttttagGACTGACC
CGCTCAAGgtaaca ... intron 2 ... ttgtagGAAGCTTA

Figure 4.2 Nucleotide and deduced translation product sequences of human C6orf37 cDNA. The sequences are numbered starting at the predicted transcription start site (TSS; boldface nucleotide) and initiation codon of the gene and protein, respectively. A putative TATA box in the proximal promoter region is underlined with a solid line. The terminal stop codon of the amino acid sequence is indicated by an asterisk. The in-frame stop codon in the 5'-UTR is underlined with a wavy line. Conserved potential phosphorylation sites are highlighted in black. RNA destabilization elements in the 3'-UTR are outlined with boxes. A consensus polyadenylation signal is underlined with a double solid line. The sequence of human *C6orf37* has been deposited in GenBank under accession number AF350451.

```
                                   ↱ +1 TSS
  -30  GAAGCCCGGTATAAAAAAAGTACTGAAGACATTTTCCCCGCACAACTGCTAAAGCTCCAGAGACACGAGCGTGTGTGGCA    50

   51  GCAAGAGCCGCCAGTTCGGGACCACCGCAGCTGGGGTGGCAGCGGCGCAGGAGGGGTCGCGGGGAGGGAGTGGTGAGCGC   130

  131  AGGCGGCAGGGGTCTGGGAAAGACGAAGTCGCTATTTGCTGTCTGAGCGCGCTCGCAGCTCCTGGAAGTGTTGCCGCCTC   210

  211  TCGGTTTGTCTCGCCTGCTGCGCTCCTAGAAGGGCGGCCGCCTCCAGGACTGACCAGGGCCAAGTGGCGCTCGGCGGGCA   290

  291  CTAC ATG GCG GAG GGT GAA GGG TAC TTC GCC ATG TCT GAG GAC GAG CTG GCC TGC AGC CCC   351
    1        M   A   E   G   E   G   Y   F   A   M   S   E   D   E   L   A   C   S   P    19

  352  TAC ATC CCC CTA GGC GGC GAC TTC GGC GGC GGC GAC TTC GGC GGC GGC GAC TTC GGC GGC   411
   20   Y   I   P   L   G   G   D   F   G   G   G   D   F   G   G   G   D   F   G   G    39

  412  GGC GGC ACG TTC GGT GGG CAT TGC TTG GAC TAT TGC GAA AGC CCT ACG GCG CAC TGC AAT   471
   40   G   G   T   F   G   G   H   C   L   D   Y   C   E   S   P   T   A   H   C   N    59

  472  GTG CTG AAC TGG GAG CAA GTG CAG CGG CTG GAC GGC ATC CTG AGC GAG ACC ATT CCG ATT   531
   60   V   L   N   W   E   Q   V   Q   R   L   D   G   I   L   S   E   T   I   P   I    79

  532  CAC GGG CGC GGC AAC TTC CCC ACG CTC GAG CTG CAG CCG AGC CTG ATC GTG AAG GTG GTG   591
   80   H   G   R   G   N   F   P   T   L   E   L   Q   P   S   L   I   V   K   V   V    99

  592  CGG CGG CGC CTG GCC GAG AAG CGC ATT GGC GTC CGC GAC GTG CGC CTC AAC GGC TCG GCA   651
  100   R   R   R   L   A   E   K   R   I   G   V   R   D   V   R   L   N   G   S   A   119

  652  GCC AGC CAT GTC CTG CAC CAG GAC AGC GGC CTG GGC TAC AAG GAC CTG GAC CTC ATC TTC   711
  120   A   S   H   V   L   H   Q   D   S   G   L   G   Y   K   D   L   D   L   I   F   139

  712  TCG GCC GAC CTG CGC GGG GAA GGG GAG TTT CAG ACT GTG AAG GAC GTC GTG CTG GAC TGC   771
  140   S   A   D   L   R   G   E   G   E   F   Q   T   V   K   D   V   V   L   D   C   159

  772  CTG TTG GAC TTC TTA CCC GAG GGG GTG AAC AAA GAG AAG ATC ACA CCA CTC ACG CTC AAG   831
  160   L   L   D   F   L   P   E   G   V   N   K   E   K   I   T   P   L   T   L   K   179

  832  GAA GCT TAT GTG CAG AAA ATG GTT AAA GTG TGC AAT GAC TCT GAC CGA TGG AGT CTT ATA   891
  180   E   A   Y   V   Q   K   M   V   K   V   C   N   D   S   D   R   W   S   L   I   199

  892  TCC CTG TCA AAC AAC AGT GGC AAA AAT GTG GAA CTG AAA TTT GTG GAT TCC CTC CGG AGG   951
  200   S   L   S   N   N   S   G   K   N   V   E   L   K   F   V   D   S   L   R   R   219

  952  CAG TTT GAA TTC AGT GTA GAT TCT TTT CAA ATC AAA TTA GAC TCT CTT CTG CTC TTT TAT  1011
  220   Q   F   E   F   S   V   D   S   F   Q   I   K   L   D   S   L   L   L   F   Y   239

 1012  GAA TGT TCA GAG AAC CCA ATG ACT GAG ACA TTT CAC CCC ACA ATA ATC GGG GAG  AGC GTC 1071
  240   E   C   S   E   N   P   M   T   E   T   F   H   P   T   I   I   G   E    S   V  259

 1072  TAT GGC GAT TTC CAG GAA GCC TTT GAT CAC CTT TGT AAC AAG ATC ATT GCC ACC AGG AAC  1131
  260   Y   G   D   F   Q   E   A   F   D   H   L   C   N   K   I   I   A   T   R   N   279

 1132  CCA GAG GAA ATC CGA GGG GGA GGC CTG CTT AAG TAC TGC AAC CTC TTG GTG AGG GGC TTT  1191
  280   P   E   E   I   R   G   G   G   L   L   K   Y   C   N   L   L   V   R   G   F   299

 1192  AGG CCC GCC TCT GAT GAA ATC AAG ACC CTT CAA AGG TAT ATG TGT TCC AGG TTT TTC ATC  1251
  300   R   P   A   S   D   E   I   K   T   L   Q   R   Y   M   C   S   R   F   F   I   319

 1252  GAC TTC TCA GAC ATT GGA GAG CAG CAG AGA AAA CTG GAG TCC TAT TTG CAG AAC CAC TTT  1311
  320   D   F   S   D   I   G   E   Q   Q   R   K   L   E   S   Y   L   Q   N   H   F   339

 1312  GTG GGA TTG GAA GAC CGC AAG TAT GAG TAT CTC ATG ACC CTT CAT GGA GTG GTA AAT GAG  1371
  340   V   G   L   E   D   R   K   Y   E   Y   L   M   T   L   H   G   V   V   N   E   359

 1372  AGC ACA GTG TGC CTG ATG GGA CAT GAA AGA AGA CAG ACT TTA AAC CTT ATC ACC ATG CTG  1431
  360   S   T   V   C   L   M   G   H   E   R   R   Q   T   L   N   L   I   T   M   L   379

 1432  GCT ATC CGG GTG TTA GCT GAC CAA AAT GTC ATT CCT AAT GTG GCT AAT GTC ACT TGC TAT  1491
  380   A   I   R   V   L   A   D   Q   N   V   I   P   N   V   A   N   V   T   C   Y   399

 1492  TAC CAG CCA GCC CCC TAT GTA GCA GAT GCC AAC TTT AGC AAT TAC TAC ATT GCA CAG GTT  1551
  400   Y   Q   P   A   P   Y   V   A   D   A   N   F   S   N   Y   Y   I   A   Q   V   419

 1552  CAG CCA GTA TTC ACG TGC CAG CAA CAG ACC TAC TCC ACT TGG CTA CCC TGC AAT TAA GAA  1611
  420   Q   P   V   F   T   C   Q   Q   Q   T   Y   S   T   W   L   P   C   N   *        437
```

```
1612  TCATTTAAAAATGTCCTGTGGGGAAGCCATTTCAGACAAGACAGGAGAGAAAAAAAAAAAAAAAGAAAAAAAAAAAGAGTGA  1691
1692  TCCAGCCCTTATTAGGGATGTGTTTTGTGCAATGATGATATGCTCCTGGTTTTAAGTTTGGCAAAGCTTATGTATCTTTT  1771
1772  AATAGATGTGGGAGCATGATCTCGAAAGGATCCTTTTCCCTTCTCTTATTCTCCTACCCAATTGGATTCTATCCTGCAAA  1851
1852  AAAAGAGAGACCTGTCATTAGAAGCAACCAGGTTCTCCTGATACAAGAGAAGAAATGTGTGATGACAATATGGGTTTGCT  1931
1932  GTATCTGCTCCCATAGCTTTGCCATAGGAAAAAAAAAAAGTGGAAAGTTTCTTTTAAGATGGAATTCATAAAAGGGAAAAT  2011
2012  ACGGAGGAAAAAAGGTCTCACTCCAACTTGTGAATCAGTTTAGGAGTTCAGATATTAATAGTAACAATACAGGAAAAAGG  2091
2092  GGAACTCCAACGTTGGGATTACTGTCTGAGGCTTGTAGCAAGTGCTTTCTGTGGAATGATCTTGTTTTGCTAACAAACGG  2171
2172  CTTGCTCCAAATGAACAGTAGTAGGTTGGTGCAGTTCTCGTAACAATCAGCAGAACTTATGATGACACAATCCATTAATT  2251
2252  CCAGCTGCGTGCATAGATCACATTTTTAAAATGTAAAAATGCAAGCAAAAACAGCTGTAACAAAGAAAGTGTGCTCAAGG  2331
2332  ACCAAAGATTTAACAGATAAAAATACCCAATTAGAAGAGATATAGTAGACTATATGAAGAGAGATTATATTTGTTACACA  2411
2412  CCAATATACATCAAAGTGCCTGTTGCCTTCTGAAAATTTGAAGTGGCAAAATTATTTTATGGTTTAATGATTATTTTATT  2491
2492  TTATCAGGGACTGCCTCAAGAAGAAAATAACATAAGCTTGTGAATGGTGGAGAAAATGCCCTATTTTTTCTTGCAAATAC  2571
2572  TTGTATAAAGTTAACATTTGTTGATCTGATATTATCATAGGTACATGTGTATGTGTGTATAAATTATATGTGTGTGTGTA  2651
2652  TATATACATTTTATATATACATTTTATATGTATATATACACAGTAGATTGACTATGATCTAGAATAATGTCTCAAATAGG  2731
2732  AAATGTTTAAATACTGTGTGTTTTTATGTTTTCAACAGGATAACATGAGACGTGGGCATATTGCAATGATGAATTAAATC  2811
2812  CACATCTAAAAAAATTAAATGAAGGAGGGAACCAAGTAATATATTTCATAGGAAGAGCAGAAATTATACTGTTTTAGTGG  2891
2892  GATTTTTTTTTCTTTTTTTTTTTTTTCTTTGGTGAGCCATAAAATTCCACAAATGGGAGAATATTTGTTTGGCAGAGCACT  2971
2972  CTTTTTTATATTGAACTGCCATTTTGACAGTTGGAACCCATTTATTAAAAAAAAAATTGCATTCCTCTATGATGTTTAAT  3051
3052  CTAGTGGATCATGGATCAGTAATAGGCTACTTAAATCCCTGACTGCTAAAAAGGATTTCCGGTGATCTAAACACTACTTG  3131
3132  CTAATGTTTAAATGAATTTTAATGAATGCATTCTGCATTTCTGGACCACTAGAATTTAGTAATGTGAAATGACCCTTTTT  3211
3212  ACAGAATATTTGCACAATTGCTTAAAATTTATATATGAGATATATATTATATATAACATTTTATAAATCATGTCAATATG  3291
3292  AAACATCTTTGATCTGGTTGTCACACTGCATTTAAATATTTAGTACTGTACTTTAAATCGCTTTCCATTAAATCAAATCC  3371
3372  AACTTTATTTTCTTTCTTACAAAAATACCAGTTATACCTTTGTGAAATGAACTGGCATTACTATTTCAGTTCAATAACAG  3451
3452  CTAATCCTAAAACCACCCTTTCTCCTAGCCAGTAGTTCCTCTAGATACTGGTCTCTGAAAATGCATTTGTTAAAAACAAA  3531
3532  ACAAAACTAACACATAAGAACCTTCCCTTTGTGTTGTGAAACAACCACATAATCTCCACAACCTTAGTGGATGACTGCTT  3611
3612  GCTATGATAATTCCTCGAAGACCCAATTAGAAGATTTTCATCATCAGTTAAAGAGAGACCACGGGAGAAAAAAATATCCT  3691
3692  CCTGTTGGCAGTATAATTTGTTTGTTTGTTTATCTAGGGATCCTCAGATGCTTAGTGCTAGGTTAATCCAGGTTAATCCG  3771
3772  TCTGGACTACCTTTTGTGCATCTTTCTTTGAAGCCTTAATGGGAACCTGATGGGTTTGCTGTAGCAGCTTCCTTGTGAAT  3851
3852  TCTGTCAGAGCTGCAACAGCCGCTGCACTGCCACTCAGTTTTCTAAGGAACTCCTCCTACTACCATCTTGGCTCAGTCTC  3931
3932  CCTCACTTAAGCCCTGGGTTTGAAAAATTAATTGCAACTTCCCAGGAAACATTGTTCAGTTTGCAGATTAAGCCTGGCAC  4011
4012  TCACCTATCAGAAACCAGAGCTCCGCCTGCTTAGTTGTTTCAAAGTTTTCTGAAAGAAAACTAGGGGAGCACTTGTGAAC  4091
4092  ACAGGAGCAGCTGGTGATCTGCTTTCTTACCCTAACTCTTGACAAATGAGTCGTCTACTATTTTAAAGAGTCTGGAGGTC  4171
4172  TCTGACTCTGCCATAACAATAACCTGCTGTTAATTTATAACACAGATTTTTGTTTGGAAGAGCCTTATTTGAAATACACT  4251
4252  TTGATTTATTTTCTTAAATATTTATATTCTTTTCTTGCTTACTTCAGGGTTGGTAGCTTAGTTGGAAGTGCCAGCACCTG  4331
4332  GCACCTATTCATATAGAACAGGCTGTACTCAAGACAACTTCTAGCATTTACTTTAAGACTTATATAATTTATTTCTATTT  4411
4412  TGTGTGTACTATAGTCTTGTGCATATGTAGTTGAACACACAGTGAAATATATGTCTCTCTTTGTGGATGTGCGGCCTAAA  4491
4492  AATTTGAATGTCTGGTGAGAGAGAGCCATGTGTATAGGTCAGAGAAAAGAACAGCTCCCGACTCCCTATTAGCGCCTGTG  4571
4572  ATTTGTTTCCTTTTGTGTTTATCTGGCCTAGTGTGCTGTTTCTTTAAACCAGGAAGAAGTTTTGTCTTTTGGAGGCTCTT  4651
4652  CTCACCTGTCCAGCCTGGCATGTCAGAGAACACATAGCCTGTGACAATGCCGTTTTTAAAGGTTTACTTAATTTGCAGTA  4731
4732  AATCCAGCTGCCTCAAGAACTCCTACACCAAGATGGACATTTCCTTTCCAGAAATGGGATCAAGTATCTGCTCACTTTGG  4811
4812  TATTGGATGGACTAATAATGTAGCTCCAAAAATGCAAGGATGGAAGAATATGTGTAATCCAAACCAAGGAAGGAAATGAA  4891
4892  AAGTGAACGTACTGTTTTTACCACCCCTTTCTGTTTGCTTATTGTTGGTTGCTTCACTGTGCATAAAGTTGTTTTCAATG  4971
4972  CAACGCTTGTTAAATAAATATTGTGAACTATTTTGTAAATGAAATGTATTATGTTGAAAGCTGTCAGTTCAAAAATAAGC  5051
5052  TTTTTTGTTGTTGTTGAAGATGAAGTGTGTTAGGTGAAACCAAAAAGCCAAAAAAAGTAATTTCATATATAGCATCTATT  5131
5132  TGAATATAATCTTTCTTTAAAATTTCTTTTAGCATAGCATTTTCAGTGCTAAGAAAGAATCTCTATGTTATATTTTTGTTA  5211
5212  AAATAATGGCTTTCTAACAAAGCAAATGGTAAAGTACAAAGTTGGAAGATGTCAAGTTAACGAGACTTGCTGCAAAGCCT  5291
5292  TGCAGAACGGAGGAGGCTCGCCTGCTGGCTGTCTCTCCCTCCAACCTCTCTACAATCATGCCTGCTTTGAGGTGTTCTGT  5371
5372  TGCAGCAAGCTGCACCTTGGGTCACTCTTTTGGAATATTTTGACTATAGGCTGCGTCACAGGCAGAAAAGGAGTTGATGG  5451
5452  AAAATGGACTAAAAAACTGACATGTTTGAATCAGTGCTAGAGGGAACAGATTGTGAATTTTGTTTACAGCATCCAATATT  5531
5532  TGGATTTTTTTGTAAATAAAAAAGTTATTTTTTTCTATTG                                          5571
```

Figure 4.3 Sequence alignment of the human C6orf37 protein and homologues identified in other species. The amino acid sequences of the *H. sapiens* (GenBank accession number AF350451), *R. norvegicus* (deduced from translation of overlapping ESTs), *M. musculus* (GenBank accession number BAB29768), *D. melanogaster* (GenBank accession number AAF59198), and *C. elegans* (GenBank accession number T15196) homologues were aligned using the Clustal W program (version 1.81). Hyphens represent gaps introduced to maintain the alignment. The amino acid residues are numbered to the left and right of the sequences. Identical residues among the sequences in the alignment are shown on a black background, and conservative amino acid substitutions are shaded in grey. Residues identical in all five homologues are marked with asterisks. The conserved potential phosphorylation sites predicted are indicated by arrows.



H. sapiens
R. norvegicus
M. musculus
D. melanogaster
C. elegans

```
H. sapiens       389  -------------------------------------------------------VIPNVA  394
R. norvegicus    379  -------------------------------------------------------VIPNVA  384
M. musculus      322  -------------------------------------------------------IIPSAT  327
D. melanogaster  415  QQQQQQQRQHQFLDPAQQQQQQQQTQTQQQQAQQATTLTLVPQAPPTGQAQTIYVQQAAP  474
C. elegans       435  ------------------------------------------------------------  435

H. sapiens       395  NVTCYYQPAPYVADANFSNYYIAQVQPVFTCQQQTYSTWLPCN-----------------  437
R. norvegicus    385  NVTCYYQPAPYVADANFSNYYIAQVQPVFTCQQQTYSTWLPCN-----------------  427
M. musculus      328  NVTCYYQPAPYVEDGNFNNYYIAHPP---ITYSTPYPTWLPCN-----------------  367
D. melanogaster  475  AAVCCTSSADQTETGAPQTIQIQAGPPGLIYANGVYYAPEQEKEQQHHQHEQQQFQQQQW  534
C. elegans       436  -MQCYTPPNFSSGSGTPPRTTLFYLPPDAATWIPVI------------------------  470
                         *

H. sapiens            ------------------------------------------------------------
R. norvegicus         ------------------------------------------------------------
M. musculus           ------------------------------------------------------------
D. melanogaster  535  SEDALGATLLAETETEPPLATASIAPPPPSSQPPSSPSTTAPRTQVQNQTRSTTINSSNR  594
C. elegans            ------------------------------------------------------------

H. sapiens            ------------------------------------------------------------
R. norvegicus         ------------------------------------------------------------
M. musculus           ------------------------------------------------------------
D. melanogaster  595  KKKNKAKTQSKQQTQTATEAESNCNEKPTTATMTRTTTKKSSRIKQMAAAAAATEACAAK  654
C. elegans            ------------------------------------------------------------

H. sapiens            ------------------------------------------------------------
R. norvegicus         ------------------------------------------------------------
M. musculus           ------------------------------------------------------------
D. melanogaster  655  GGAGGITGSLVLTPQQINGLPVYNQLDDGGRSAAGTYSGQFAYFYPHDGSAPSMFLLGED  714
C. elegans            ------------------------------------------------------------

H. sapiens            ------------------------
R. norvegicus         ------------------------
M. musculus           ------------------------
D. melanogaster  715  SYLNVTEQQEQHQHQHQHHNHQFV  738
C. elegans            ------------------------
```

Figure 4.4 ***In vitro*** **translation product of** ***C6orf37*** **coding sequence.** An autoradiograph of the ^{35}S-methionine-labeled product resolved by SDS-PAGE is shown. Positions of protein size markers are indicated on the left (*kDa*, kilodaltons). Image provided by Ms. Laura E. Kakuk (University of Michigan).

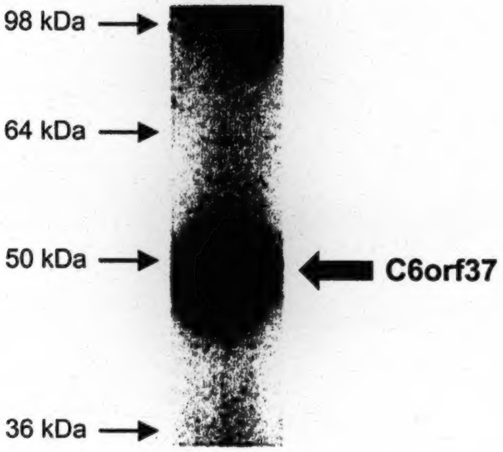
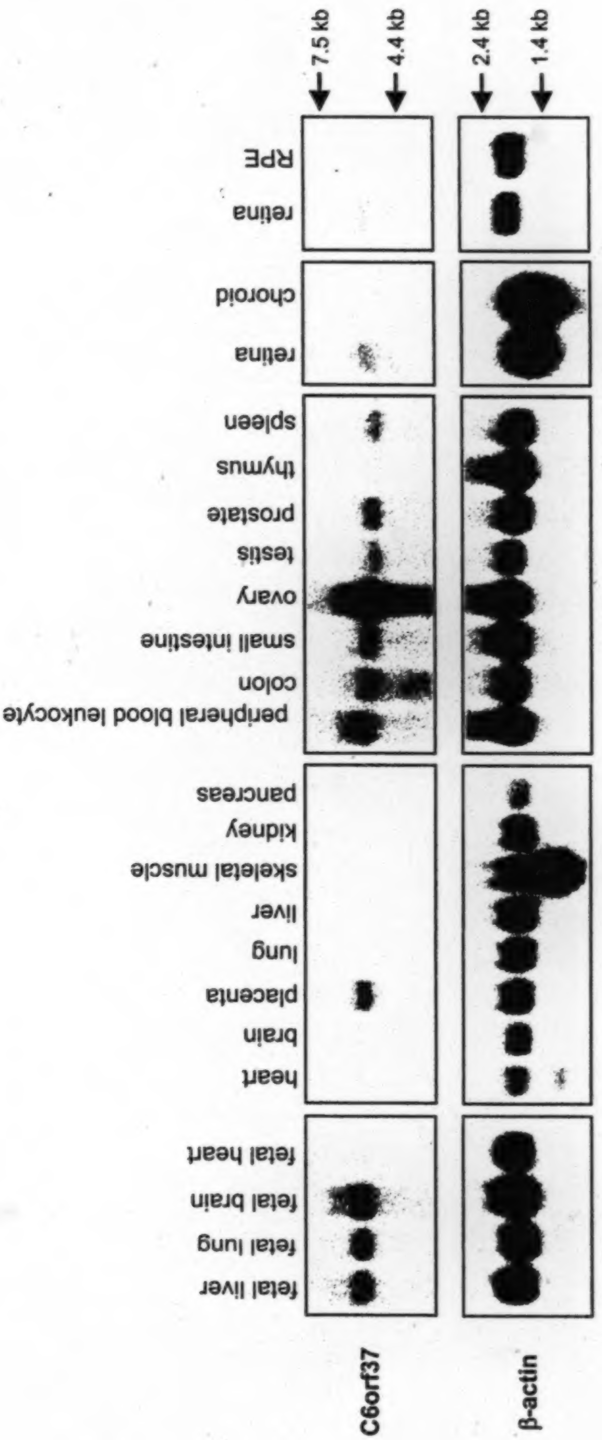

98 kDa
64 kDa
50 kDa
36 kDa
C6orf37

Figure 4.5 Expression profile of the *C6orf37* gene in human tissues. Northern blots of human RNA samples probed with C6orf37 and control (β-actin) cDNA probes are shown. Retina, choroid, and retinal pigment epithelium lanes contain 5 μg of total RNA while all other lanes contain 2 μg of poly(A)$^+$ RNA. Positions of RNA size markers are indicated on the right (*kb*, kilobases).


C6orf37
β-actin
fetal liver
fetal lung
fetal brain
fetal heart
heart
brain
placenta
lung
liver
skeletal muscle
kidney
pancreas
peripheral blood leukocyte
colon
small intestine
ovary
testis
prostate
thymus
spleen
retina
choroid
retina
RPE
7.5 kb
4.4 kb
2.4 kb
1.4 kb

Chapter 5

A 5-bp deletion in the novel retinal gene *ELOVL4* is associated with related forms of autosomal dominant macular dystrophy

A version of this chapter has been published.
Zhang K, Kniazeva M, Han M, Li W, Yu Z, Yang Z, Li Y, Metzker ML, Allikmets R, Zack DJ, Kakuk LE, Lagali PS, Wong PW, MacDonald IM, Sieving PA, Figueroa DJ, Austin CP, Gould RJ, Ayyagari R, Petrukhin K. A 5-bp deletion in *ELOVL4* is associated with two related forms of autosomal dominant macular dystrophy. *Nature Genetics* 2001;27(1):89-93.

5.1 INTRODUCTION

Stargardt-like macular dystrophy (STGD3, OMIM #600110) and autosomal dominant atrophic macular degeneration (adMD) are inherited forms of macular degeneration characterized by decreased visual acuity, bilateral atrophic lesions of the macula and central retinal pigment epithelium, and the presence of yellow fundus flecks.[1-5] Though differences in age of onset and variable expressivity have been noted in different families, these diseases bear striking clinical resemblance.[1-5]

Genetic mapping of the disease locus for autosomal dominant Stargardt-like macular dystrophy was initially performed for a five-generation family in the United States and revealed linkage of the disease-causing gene to a pericentromeric region on the long arm of human chromosome 6.[2] Subsequently, two additional families sharing common ancestry were described,[3] and the disease gene in all three independent kindreds was found to map to the STGD3 locus on chromosome 6q14. These two studies defined the minimal genomic interval harbouring the disease gene to be 18 cM and 21 cM respectively, with the former interval lying entirely within the latter.[2,3]

An adMD pedigree exhibiting phenotypic characteristics of autosomal dominant macular degeneration similar to those observed in STGD3 patients was identified, and linkage analysis placed the corresponding disease locus in an 8-cM region of chromosome 6q14 contained within the previously defined STGD3 interval.[4] More recently, a fourth STGD3 family living in Alberta, Canada was described, with a 23-cM minimal critical interval overlapping all the previously delineated STGD3 and adMD loci.[5] It has been speculated that the similar though somewhat variable diseases affecting these five families may be attributed to allelic mutations in the same gene or may arise from locus heterogeneity, i.e. mutations in separate genes residing on 6q14.[4,5]

In the genetic studies presented here, the minimum genetic region for STGD3 and adMD is limited to a 0.6-cM interval by recombination breakpoint mapping among individuals from all five families which were determined to share a common founder haplotype. Through subsequent candidate gene identification procedures and mutation screening, a single five base-pair deletion was identified within the protein-coding region of a novel retinal gene called *elongation of very long chain fatty acids 4* (*ELOVL4)* in all affected members of the STGD3 and adMD families. Initial bioinformatic analysis of the ELOVL4 protein revealed its homology to a group of yeast proteins termed elongases that function in the biosynthesis of very long chain fatty acids.[6,7] These results are therefore the first to implicate the biosynthesis of fatty acids in the pathogenesis of inherited macular degeneration.

5.2 MATERIALS AND METHODS

5.2.1 Family Material and Linkage Analysis

Each participant of five large Caucasian families (A, B, C, STGD3/Canadian (see Chapter 2), and UM:H785) underwent a detailed ophthalmologic examination including a visual acuity measurement and a dilated ophthalmoscopic examination. A diagnosis of Stargardt-like macular dystrophy in families A, B, and C was made by Dr. Kang Zhang and associates (John Hopkins University, Baltimore, MD) based on decreased visual acuity and characteristic fundus appearance. Diagnostic criteria and genetic characterization of families UM:H785 and STGD3/Canadian were described previously.[4,5] Two-point linkage analysis in the five families using the marker *D6S460* yielded a maximum LOD score of 9.66 (Family A), 4.36 (Family B), 4.42 (Family C), and 3.34 (STGD3/Canadian) at θ=0 and 3.5 at θ=0.1 in UM:H785. The number of family members used in linkage analysis and which were screened for a mutation in the *ELOVL4* gene was (total/affected individuals) 46/26 for Family A, 19/12 for Family B, 20/11 for Family C, 23/11 for STGD3/Canadian family, and 39/16 for UM:H785. Family C overlaps with a previously reported 10-generation STGD3 pedigree.[3] Family A represents a previously described 4-generation pedigree with the reported linkage to chromosome 13q34 (Zmax=4.6 at θ=0.15 with *D13S71*).[8] Additional clinical studies of children who were initially too young to establish a diagnosis and subsequent genetic linkage (Zmax=9.66 at θ=0 with *D6S460*) and haplotype analysis placed the disease gene responsible for the Stargardt-like phenotype in this family to the *STGD3* locus on human chromosome 6q14, thus eliminating the STGD2 locus for autosomal dominant Stargardt-like phenotype on 13q34.

5.2.2 *ELOVL4 cDNA Cloning*

DNA sequences of PAC/bacterial artificial chromosome (BAC) clones comprising the Chr_6ctg16 contig were downloaded from the The Sanger Center web site (http://www.sanger.ac.uk/HGP/Chr6/), and the size of the minimum genetic region was estimated to be within the 400-600-kb range. DNA sequences of 11 fragments from PAC dJ92C4 generated in Phase 1 High Throughput Genome Sequence Project (GenBank accession number AL132875) were compared with GenBank database protein entries using the BLASTX algorithm. This comparison revealed the presence of two potential exons of *ELOVL4* (exons 2 and 4). To identify additional exon(s) that may be located between exons 2 and 4, forward and reverse PCR primers from these exons were synthesized and used to PCR-amplify ELOVL4 cDNA fragments from human retina "Marathon-ready" cDNA (BD Biosciences Clontech) by Dr. Konstantin Petrukhin (Merck Research Laboratories, West Point, PA). In this RT-PCR experiment, a forward primer from exon 2 (5'-GTGTGGAAAATTGGCCTCTG-3') was used with a reverse primer from exon 4 (5'- GTCCTCCTGCAACCCAGTTA-3'). The assembly of the DNA sequence results of this PCR product revealed the presence of exon 3. The DNA sequence of exon 4 was compared with the EST database using the BLASTN algorithm in an attempt to identify additional cDNA sequences. This analysis identified a mouse skin EST (GenBank accession number AA791133) with very high degree of similarity to exon 4 of the human *ELOVL4* gene. The DNA sequence of the mouse skin EST was compared with the genomic sequence of PAC dJ92C4. Despite the differences between the mouse and human sequences caused by evolutionary divergence, this analysis was able to reveal two additional human exons within PAC dJ92C4 that we labeled exons 5 and 6.

To identify the 5'- and 3'-ends of the human ELOVL4 cDNA, rapid amplification of cDNA ends (RACE) was performed by Dr. K. Petrukhin using the "Marathon-ready" RACE template from human retina. The assembly of the DNA sequence results of these PCR products produced a complete cDNA sequence, consisting of a total of 2969 bp, including 226 bp of untranslated sequence at the 5' end (5'-UTR) followed by a 945-bp open reading frame and a 1798-bp 3'-UTR (GenBank accession number NM_022726). The DNA sequence of the full-length mouse Elovl4 cDNA (GenBank accession number AF277093) was determined after performing a round of 5'-RACE using the "Marathon-ready" RACE template from mouse testis (BD Biosciences Clontech).

Independently, a comprehensive collection of non-overlapping sequence-verified human ESTs mapping to the minimal genomic interval containing the disease locus for the STGD3 and adMD families were tested for retinal and retinal pigment epithelial expression in an effort to define novel candidate retinal disease genes, as previously described.[9] This analysis revealed the existence of two human fetal ESTs (GenBank accession numbers AA969661 and AA993222) that exhibited both retinal and RPE expression. These two ESTs shared homology with and/or overlapped several additional catalogued EST clones that comprised a single UniGene cluster, Hs.101915. The translation product of the consensus cDNA sequence was found to share homology with the *Saccharomyces cerevisiae* proteins Elo1p, Elo2p, and Elo3p (GenBank accession numbers P39540, P25358, P40319) upon BLASTP analysis.

5.2.3 Organization of the Human ELOVL4 Gene

To verify the exonic composition of the cDNA, forward and reverse PCR primers from exon 2 (5'-GTGTGGAAAATTGGCCTCTG-3') and exon 6 (5'-GTCAACAACAGTTAAGGCCCA-3') within the genomic PAC dJ92C4 clone, respectively, were used to PCR-amplify an ELOVL4 cDNA fragment from a human retina "Marathon-ready" cDNA library (BD Biosciences Clontech) by Dr. K. Petrukhin. The DNA sequence of this PCR product provided confirmation of the cDNA sequences which were assembled from *ELOVL4* exons, and corrected the sequencing errors present in the database entry for PAC dJ92C4.

The *ELOVL4* gene structure was also determined independently by comparing the cDNA sequence assembled from constituent ESTs, identified in the screen for retina- and RPE-expressed ESTs mapping to chromosome 6q12-q16, with human chromosome 6 genomic sequence found in DNA sequence databases. The organization of the *ELOVL4* gene reported here is based on alignment of the mRNA sequence (GenBank accession number NM_022726) with the human chromosome 6 contig clone NT_007299.

GenBank accession numbers for the human *ELOVL4* exon sequences are as follows: exon 1, AF279649; exon 2, AF279650; exon 3, AF279651; exon 4, AF279652; exon 5, AF279653; exon 6, AF279654.

5.2.4 *Bioinformatic Protein Sequence Analysis*

The translated ELOVL4 cDNA sequence was subjected to bioinformatic analysis using protein sequence analysis tools available through the Canadian Bioinformatics Resource (http://cbr-rbc.nrc-cnrc.gc.ca/) and the ExPASy Molecular Biology server (http://ca.expasy.org/tools/). Amino acid sequence alignments were carried out using the Clustal W program (version 1.83).[10] Hydropathy plots were generated using the GREASE program (version 2.0) (http://fasta.bioch.virginia.edu/fasta_www/grease.htm) according to the algorithm of Kyte and Doolittle,[11] and protein topology prediction was independently verified using the TMpred program (http://www.ch.embnet.org/software/TMPRED_form.html).[12]

5.2.5 *RNA Isolation and Northern Blot Analysis*

For human retina and RPE RNA analysis, ocular tissues were extracted from donor eyes derived from a 33-year old normal male, and total RNA was isolated using the Trizol reagent (Invitrogen Canada Inc., Burlington, ON) according to the manufacturer's protocol. Five micrograms of each total RNA sample was electrophoresed on a 1% agarose-formaldehyde gel and subsequently blotted onto a nylon membrane (Pall Corporation, Mississauga, ON) via capillary transfer. For the analysis of non-ocular human tissues, multi-tissue northern blots containing two micrograms of poly(A)$^+$ RNA per tissue sample were obtained commercially (BD Biosciences Clontech).

Northern blots were consecutively hybridized with probes corresponding to human ELOVL4 and β-actin coding regions. The ELOVL4 cDNA probe used was generated by total cDNA synthesis using a commercial kit (Invitrogen Canada Inc.) and five micrograms of human retina total RNA as template, followed by PCR using 100 nanograms of total cDNA template with primers 5'-GTGTGGAAAATTGGCCTCTG-3' and 5'-CATGGCTGTTTTTCCAGCTT-3' to amplify a 743-base pair cDNA fragment containing protein-encoding sequences found within exons 2 through 6. The β-actin cDNA probe was amplified by PCR from a plasmid clone with primers 5'-GTCCACACCCGCCGCCAGCTCACCAT-3' and 5'-ACTGGTCTCAAGTCAGTGTACAGGTA-3' to generate a 1698-base pair nearly full-length cDNA product. PCR cycling parameters used were as follows: 1 cycle of denaturation at 94°C for 3 minutes; 30 cycles of denaturation at 94°C for 1 minute, primer annealing at 45°C for 1 minute, primer extension at 72°C for 3 minutes; 1 cycle of extension at 72°C for 7 minutes. Probes were purified using the QIAquick Gel Extraction Kit (QIAGEN Inc., Mississauga, ON) and radiolabeled with [α^{32}P]-dCTP (Amersham Biosciences) by random primer labeling using the Multiprime DNA Labeling System (Amersham Biosciences). Hybridization of each probe to the northern blots was carried out in Hybrisol II hybridization solution (Serologicals Corporation) at 65°C for 16 hours using 10^6 counts per minute (cpm) of labeled, purified probes per milliliter of hybridization solution. The blots were then washed twice for 15 minutes in 2X SSC, once for 30 minutes in 2X SSC, 0.1% SDS, and once for 10 minutes in 0.1% SSC, 0.1% SDS all at 65°C, followed by exposure to Kodak X-Omat film at -70°C. ELOVL4 cDNA-probed blots were exposed to film for 18-48 hours.

5.2.6 ELOVL4 Mutation Analysis

Genomic DNA from STGD3/adMD patients and normal controls was amplified by PCR using oligonucleotide primers for intronic sequence spaced 30-50 bp from exon/intron junctions in the *ELOVL4* gene. The sequences of the primer pairs used for amplification of the exons are given in Table 5.1. DNA (50-100 nanograms) was amplified using *Taq* Gold polymerase (Perkin Elmer) and a modified touch-down protocol[13] in a total volume of 50 microlitres. Modified PCR reactions for exon 1 included 10% DMSO. Amplification bands were excised from 2% agarose gels, purified using QIAquick spin columns (QIAGEN Inc.), and sequenced using ABI dye-terminator sequencing kits (Perkin Elmer). Once the 5-bp deletion had been identified through exon 6 sequence analysis, the forward primer 5'-GCAATATTCCCCAAATGGATGCACTGG-3' and reverse primer 5'-TTGCGGCCGCTTAGGCTCTTTGTATGTCCGAAT-3' were used to amplify a fragment of exon 6 containing the deletion site for subsequent mutation detection by polyacrylamide gel electrophoresis. PCR reactions for amplification of these smaller exon 6 fragments were performed using a radioactive PCR procedure previously described[5] with 200 nanograms of genomic DNA as template and 0.45 units of Taq polymerase (QIAGEN Inc.) in a reaction volume of 15 microlitres. PCR products were electrophoresed on a denaturing 6% polyacrylamide gel and visualized by autoradiography.

5.3 RESULTS

5.3.1 Defining a 0.6-cM Minimal Critical Interval for STGD3/adMD

To identify the disease gene causing heritable macular dystrophy, initial studies were performed on three independent kindreds in the United States with autosomal dominant Stargardt-like macular dystrophy characterized by bilateral, symmetric, atrophic lesions with sharp borders in the macula and by the frequent presence of yellow fundus flecks. A genetic linkage analysis placed the disease gene in all three families to the known *STGD3* locus on chromosome 6q14.[2,3] The position of the *STGD3* locus was refined by meiotic breakpoint mapping and the critical region was limited to an approximately 1.8-cM interval between markers *D6S1622* and *D6S391* (Figure 5.1A). To further reduce the minimum genetic region, a five-generation family with STGD3 from Canada (see Chapter 2 for details),[5] as well as an independent five-generation family from the U.S. with adMD[4] were recruited into the study. Genetic linkage and haplotype analyses performed in both of these families were previously reported and confirmed the localization of disease loci to the STGD3 interval on chromosome 6q14.[4,5] Comparison of disease haplotypes revealed that all affected members in the five independent families with STGD3 and adMD shared a common founder haplotype between markers *D6S1625* and *D6S1644*, indicating a single, ancestral, disease-specific mutation. A recombination event in individual V-14 from family UM:H785[4] at marker *D6S460* identified the centromeric boundary of the critical interval, thus limiting the minimum genetic region to a 0.6-cM interval between markers *D6S460* and *D6S391* (Figure 5.1A).

5.3.2 Identification of the ELOVL4 Gene

To identify candidate genes in the *D6S460-D6S391* interval, DNA clones derived from the inclusive PAC/BAC contig Chr_6ctg16 assembled at the Sanger Center (http://www.sanger.ac.uk/HGP/Chr6/) were analyzed (Figure 5.1B). BLASTN analysis of DNA sequences from PAC clones covering the region revealed the presence of two candidate genes, chromosome 6 open reading frame 7 (*C6orf7*) and the gene encoding the E1 β-subunit of branched chain α-ketoacid dehydrogenase (*BCKDH*) (Figure 5.1C), in which amino acid-changing or splice-site mutations were not found. In an attempt to define additional novel candidate genes located within the minimal critical region, human

ESTs that were known to map to chromosome 6q14 were screened for retinal or RPE expression. This analysis resulted in the identification of two retina- and RPE-expressed ESTs (Figure 5.1D) that were part of a larger EST cluster whose consensus sequence, when translated, showed homology with the members of the yeast elongase (ELO) family known to be involved in the biosynthesis of fatty acids.[6,7] To account for the possibility that some retinal genes might not be represented in EST databases, the search for candidate genes was broadened by comparing the DNA sequence obtained from the PAC clones with GenBank protein entries using a BLASTX algorithm. This comparison revealed two potential exons in PAC dJ92C4 whose DNA sequences were also found to share homology with the yeast ELO proteins upon translation. *ELOVL1, ELOVL2,* and *ELOVL3* are yet to be cloned human orthologues of the murine lipid elongases *Elovl1, Elovl2,* and *Elovl3*, formerly known as *Cig30*, *Ssc1*, and *Ssc2*, respectively.[14,15] Therefore the new human retinal gene identified was named *ELOVL4* (for elongation of very long chain fatty acids 4). Comparison of the cDNA sequence with genomic DNA enabled the characterization of the organization of the human *ELOVL4* gene, which was found to consist of six exons spanning approximately 33 kilobases on chromosome 6 (Figure 5.1E).

5.3.3 Characterization of ELOVL4 Gene Products

Translation and BLASTP analysis of the ELOVL4 cDNA revealed that it encodes a putative protein of 314 amino acids (Figure 5.2) with 93% amino acid identity to the mouse orthologue identified and approximately 25% identity to members of the *ELO* gene family in yeast, which encode components of the membrane-bound fatty acid elongation system.[6,7] Bioinformatic analysis of the human ELOVL4 protein sequence revealed the presence of all three features typical for members of the ELO family: a single HXXHH 'histidine box' motif characteristic of diiron-oxo cluster proteins involved in redox reactions[16,17] (Figure 5.2); a strong signal (dilysine motif with the lysines at –3 and –5 positions relative to the carboxyl terminus) shown to be responsible for the retention/retrieval of transmembrane proteins in the endoplasmic reticulum[18-20] (Figure 5.2), the known site of biosynthesis of fatty acids with very long chains[21]; and a hydropathy plot that predicts 5 transmembrane segments (Figure 5.3).

Northern blot hybridization of an ELOVL4 cDNA probe to human retina RNA showed abundant expression of an approximately 3.0-kb transcript corresponding to the expected full length mRNA (Figure 5.4). This transcript was also found in human brain, thymus and fetal heart tissue, indicating that expression of the *ELOVL4* gene is not retina-specific but is restricted to particular tissues. A less abundant 2.4-kb transcript was also detected, which is consistent with the size of a second transcript possibly arising from usage of an alternate polyadenylation signal found within the 3'-UTR of the *ELOVL4* gene (Figure 5.1E). This transcript was observed for retina, lung and thymus RNA samples. Additional bands of various sizes were also apparent on the blots, suggesting that additional transcripts that share significant homology to the cDNA probe used are expressed in some human tissues.

5.3.4 Identification of an ELOVL4 Mutation in STGD3/adMD Patients

The absence of amino acid-changing or splice-site mutations from *C6orf7* and *BCKDH* indicated that *ELOVL4* was the most likely candidate for STGD3 and adMD. Mutation analysis of the new candidate gene was performed in four STGD3 kindreds and one family with adMD. Flanking intronic primers (Table 5.1) were used to amplify and sequence all 6 exons of the candidate gene from several unaffected and heterozygous affected individuals. DNA sequence analysis revealed the presence of a 5-base pair deletion in exon 6 starting at position 790 of the ELOVL4 cDNA relative to the translation start site, designated 790-794delAACTT, in affected individuals (Figure 5.5A). The deletion results in a frameshift, loss of 51 amino acids at the carboxyl terminus of the wild-type protein, and synthesis of an aberrant peptide from amino acids 264 to 271 followed by a premature stop codon. The presence of 790-794delAACTT was assessed in all members of the five families by polyacrylamide gel electrophoresis of exon 6 PCR products (Figure 5.5B) followed by direct sequencing. The deletion was present in all affected individuals who inherited the founder haplotype (a total of 76 chromosomes in 5 families) and was absent in unaffected relatives (Figure 5.5). Consistent with the fact that 790-794delAACTT is not a common polymorphism, this allele was not found in any of the normal chromosomes from 50 unrelated individuals of predominantly European descent. Moreover, this mutation was not found in 243 controls and 245 AMD individuals comprising a previously described clinical sample.[22]

5.4 DISCUSSION

The ELO gene family was initially identified in *Saccharomyces cerevisiae* as encoding enzymatic components of medium- and long-chain fatty acid elongation pathways that yield products of increased carbon chain length.[6,7] Several other members of this gene family were subsequently found in rodents, where they were similarly shown to be necessary for fatty acid chain elongation.[14,15] Additional elongases have since been cloned from a variety of species including plants,[23-26] the nematode worm *Caenorhabditis elegans*,[27,28] and also humans.[29] Engineered mutations in the yeast ELO genes markedly alters cellular fatty acid composition and leads to pleiotropic effects involving a number of membrane functions.[7,30] Ablation of ELO gene expression by mutation or RNA interference in *C. elegans* results in fatty acid composition imbalances and multiple phenotypic defects in growth, reproduction, and physiological rhythms.[31,32] Recently described *Elovl3* knockout mice exhibit abnormalities in the lipidation and resultant function of skin and hair.[33] While other examples of ELO gene disruption have not been documented, the results of the yeast, nematode and mouse studies as well as the biological importance of very long chain fatty acids in animals[34] suggest that mutations in human ELO genes would have deleterious effects.

Most ELO genes have been identified based on sequence homology to previously isolated gene family members in the same or other species. *ELOVL4* represents the first fatty acid elongase gene identified through a combined positional-candidate cloning strategy towards isolation of a disease gene. Bioinformatic analysis revealed structural and topological similarities between the protein encoded by this novel gene and previously defined ELO homologues, providing clues to its function that further support its characterization as a plausible candidate for retinal dysfunction. Based on protein sequence analysis and gene expression observed in the retina, it seems reasonable to suggest that ELOVL4 represents a component of the fatty acid elongation system residing in the endoplasmic reticulum (ER) within human retinal cells. The process of fatty acid elongation in mammalian cells involves a series of condensation, reduction, and dehydration reactions using substrates derived from either *de novo* fatty acid synthesis pathways or dietary sources, and is catalyzed by a closely coordinated enzyme system.[20,35] Biosynthesis of docosahexaenoic acid (DHA), which contains 22 carbon atoms (C22) and represents approximately 50% of the fatty acids in retinal photoreceptor cell outer segment phospholipids,[36] requires dietary consumption of the essential α-linolenic acid and a series of three elongation steps alternating with chain desaturation events.[37,38] ELOVL4 may be involved in one of these enzymatic reactions required for DHA biosynthesis, and/or perhaps those yielding other fatty acid elongation product(s) critical for retinal function. The HXXHH histidine box motif found in the ELOVL4 amino acid sequence is characteristic of diiron-oxo proteins in which the histidine residues assist in the coordination of a Fe-O-Fe cluster at the active site, which acts as an electron acceptor in reduction reactions.[16,17] The identification of this motif in a condensing enzyme identified in mouse that shares homology with ELOVL4,[39] coupled with the observation that protein database searches fail to detect any sequence homology between ELOVL4 and either of the recently identified mammalian reductase enzymes,[40] nor any of the numerous HXXHH-containing desaturase genes identified that participate alongside elongases in polyunsaturated fatty acid synthesis,[41] suggests that ELOVL4 may represent an enzyme that catalyzes the first and rate-limiting step in the fatty acid elongation pathway. Confirmation of this hypothesis requires characterization of the specific enzyme activity and substrate specificity of the ELOVL4 protein.

Retinal expression of *ELOVL4* might therefore be predicted, given the unique fatty acid composition of photoreceptor cell membranes,[36] the lipid-dependent activities of constituent membrane proteins that mediate visual processes,[42,43] and the importance of dietary long-chain fatty acids, i.e. elongation reaction precursors, to the development of normal vision.[44] Essential roles of very long chain fatty acid synthesis in other tissues may also explain the extra-retinal *ELOVL4* gene expression observed. For example, C20 and C22 metabolites of essential fatty acids play a major role in brain development and function,[45] and various long-chain polyunsaturated fatty acids influence the development and regulation of immune function as well as the severity of autoimmune diseases that target immune cells of the thymus.[46] Studies have shown that fetal primates and preterm human infants are capable of synthesizing DHA from parental fatty acid precursors,[47] implying the existence of fatty acid elongases in fetal tissues, which is supported by our observations.

Despite these associations and the observation of both ocular and non-ocular *ELOVL4* gene expression, STGD3/adMD disease pathology appears to be limited to the retina. Based on both EST tissue source data and northern analysis, it is clear that *ELOVL4* mRNA expression is not retina-specific, however the neural retina appears to express a unique collection of *ELOVL4* transcripts detectable by hybridization when compared with other human ocular and non-ocular tissues. In mice, ELOVL4 mRNA levels are dramatically higher in the retina compared to any other tissue examined, including brain.[48] Thus differential rather than tissue-restricted expression may impart some level of gene regulation leading to functional differences between sites of expression and may account for the retina-specific defects observed in patients with *ELOVL4* mutations. Alternatively, there may exist redundant pathways in other tissues that compensate for the mutation, or perhaps the retina may be less tolerant to a reduction in functional ELOVL4 levels compared to other tissue types. In support of this theory, mouse *ELOVL4* gene expression in the brain has been shown to peak during development, while retinal expression persists at high levels throughout adulthood.[48] Another possible mechanism for retina-specific dysfunction resulting from *ELOVL4* mutation is the involvement of tissue-specific factors in ELOVL4-mediated processes. Thus a particular mutation in *ELOVL4* may disrupt association of the protein with a critical retina-specific component essential for fatty acid elongation in this tissue, but which is not present or essential for function in other tissues. This has been observed for the retinitis pigmentosa GTPase regulator (*RPGR*) gene, a ubiquitously expressed gene in which mutations abolish binding of the protein to photoreceptor-specific interacting proteins[49] and which lead to exclusively retinal degenerative phenotypes. A growing number of widely expressed genes leading to retina-specific phenotypes upon mutation have been documented.[50]

Combined haplotype analysis performed for the STGD3/adMD families significantly reduced the size of the disease gene-containing interval to 0.6 cM, however the minimum genetic region was not limited to an interval entirely contained within the *ELOVL4* gene. Therefore, the possibility of a disease-causing mutation residing in an adjacent gene within the 0.6-cM interval cannot be excluded. However, linkage disequilibrium between the STGD3/adMD phenotypes and the 790-794delAACTT variant (absent from >1,100 non-STGD3/adMD alleles), *ELOVL4* expression in the retina, the severe nature of the mutation (deletion of the carboxy-terminal portion of the protein that includes a putative targeting signal required for retention/retrieval of transmembrane proteins in the ER[18-20]), and a high degree of sequence conservation between human and mouse homologues in the region of the mutation and the subsequently deleted amino acids indicate that the 5-bp deletion is the disease-causing mutation in this family. The absence of amino acid-changing or splice-site mutations from two other closely linked potential candidate genes found in the *D6S400-D6S391* interval, *C6orf7* and *BCKDH*, further supports this argument.

The 5-bp deletion was found to associate with the disease phenotype in all 38 of the affected members of the STGD3/adMD families assessed, as predicted by genealogical analysis revealing common ancestry. This has recently been extended to include additional related kindreds,[51,52] bringing the current number of branches within the entire family to 31, with 8 branches thus far known to contain macular dystrophy-affected individuals. The utility of large pedigrees towards the isolation of disease genes is clearly demonstrated in this study, and also greatly impacts genetic disease susceptibility assessment and genetic counseling for known and newly identified family members.

Disease pathogenesis in the STGD3/adMD patients is thought to be due to the loss of the ER-targeting signal at the carboxyl terminus of the ELOVL4 protein resulting from the frame shift and creation of a premature stop codon in mutant transcripts. As a result, the protein produced would be prevented from reaching its correct subcellular destination, the site of fatty acid elongation, thereby rendering it nonfunctional in the lipid biosynthetic process. In this scenario, the reduction in the amount of active ELOVL4 enzyme due to the 790-794delAACTT mutation would likely result in a deficiency in very long chain fatty acids, and a corresponding increase in precursor molecules, altering the cellular lipid composition and thus having particularly deleterious consequences in photoreceptor cells of the retina. Studies have shown that the optimal activity of rhodopsin depends on lipid-protein interactions in rod outer cell segments.[42,53] In addition, the DHA content of photoreceptor cells affects retinoid recycling and transport between the retina and RPE.[54] The *ABCA4* gene encodes a retina-specific adenosine triphosphate (ATP)-binding cassette (ABC) transporter (ABCR), and is mutated in autosomal recessive Stargardt disease (STGD1).[55] ABCR functions as an outwardly directed flippase for removal of N-retinylidene-phosphatidylethanolamine from photoreceptor outer segments,[56,57] and its activity is also strongly dependent on the surrounding lipid environment.[43] The role of *ELOVL4* in lipid biosynthesis might therefore help to explain the phenotypic similarities between recessive and dominant forms of Stargardt macular dystrophy.

The identification of a dominant mutation in *ELOVL4* leading to disease indicates that pathogenesis is due to either haploinsufficiency or a dominant negative effect. Although the majority of inborn errors of metabolism due to an enzyme deficiency are the result of recessive mutations,[58] it is plausible that the reduced level of ELOVL4 protein activity, due to the presence of only a single functional allele, rather than its complete absence may be sufficient to yield a retinal dystrophic phenotype, particularly given the unique composition and role of lipids in photoreceptor cells. In addition, the accumulation of fatty acid precursors that would otherwise act as substrates for the ELOVL4 enzyme could also lead to altered lipid profiles within retinal cells and possibly phototoxicity due to a build up of lipid peroxidation products, commonly thought to cause retinal cell degeneration.[59,60] A dominant negative effect may arise from deleterious activities of the truncated protein encoded by the mutant *ELOVL4* allele. It is believed that the enzymes of the fatty acid elongation pathway act in a closely coordinated manner,[35] possibly as a multi-subunit complex, within the ER membrane. Thus mutation of ELOVL4 may render other members of the putative complex inactive or less active, for example by causing their destabilization within the membrane due to the loss of stabilizing protein-protein contacts between subunits, reduced enzymatic activity and/or reaction rates due to the lack of appropriate substrates normally provided by upstream ELOVL4-catalyzed reactions, or failure of essential protein components dependent on ELOVL4 interactions to be recruited to the fatty acid elongation enzyme complex.

Despite harbouring the same founder mutation, different branches of the STGD3/adMD extended family exhibit variability in some clinical features and age of onset of retinal pathology. This may be due to the effects of environmental factors and/or genetic modifier loci that act in the same or a parallel biological pathway as *ELOVL4* that influence disease progression. This has also been proposed for a second mutation more recently identified in the *ELOVL4* gene in a pedigree unrelated to that described here and which causes diverse macular dystrophy phenotypes among affected family members.[61] A candidate modifier gene is *ABCA4*; the similarity of STGD1 and STGD3/adMD phenotypes, the membrane lipid environment-dependent function of the ABCR protein,[42] and the more severe clinical presentation of a STDG3 patient who is also a heterozygous carrier of a variant ABCA4 allele[62] all suggest functional interaction between the *ELOVL4* and *ABCA4* genes. Thus the clinical variability observed in STGD3/adMD may be due to multiple genetic mutations giving rise to polygenic phenotypes. Such a phenomenon has previously been observed for individuals with digenic retinitis pigmentosa due to mutation in both the *ROM1* and *peripherin/RDS* genes.[63]

The observation that STGD3/adMD patients primarily exhibit defects in the foveal-macular area of the retina might be explained by regional differences in *ELOVL4* gene expression within the retina. Alternatively, the differing lipid composition between the macula and peripheral retina[64] as well as the increased susceptibility to oxidative stress due to lipid peroxidation in the macula as compared to the peripheral retina[65] may suggest possible biochemical mechanisms for macular dysfunction resulting from *ELOVL4* mutation. In addition, a greater density of RPE cells in the central retina compared to that in peripheral regions has been observed[66]; thus if aberrant *ELOVL4* function leads to a primary defect in the RPE causing a secondary effect on photoreceptor cell survival, as has been proposed for the mutant ABCR protein,[67] it is reasonable to predict that a macular phenotype would predominate in both recessive and dominant forms of Stargardt disease, as is observed. Understanding the pathogenesis of STGD3/adMD may therefore provide insight into the etiology of other maculopathies, including age-related macular degeneration.

Continued study of the *ELOVL4* gene will increase our knowledge of its cellular role in the normal retina as well as the functional consequences of mutations leading to retinal defects. Understanding in greater detail the biology of this novel member of a gene family associated with lipid metabolism and now also proper visual function has implications for both the diagnosis and potential treatment of macular disease.

5.5 ACKNOWLEDGEMENTS

The patients with STGD3 and adMD and members of their families are thanked for their participation in this study. This work was supported by NIH EY00401 and the Grant Ritter Fund (KZ), the Foundation Fighting Blindness, Hunt Valley, MD (DJZ, PAS, RA), the E.A. Baker Foundation of the Canadian National Institute for the Blind, the Canadian Institutes of Health Research, the Alberta Heritage Foundation for Medical Research and the Foundation Fighting Blindness, Canada (PWW, PSL), and the Howard Hughes Medical Institute of which MH is an assistant investigator. This research was supported in part by funding provided to Johns Hopkins University under a Research Collaboration Agreement between Johns Hopkins University and Merck & Co., Inc.

5.6 REFERENCES

1. Cibis GW, Morey M, Harris DJ. (1980) Dominantly inherited macular dystrophy with flecks (Stargardt). *Arch. Ophthalmol.* 98:1785-1789.

2. Stone EM, Nichols BE, Kimura AE, Weingeist TA, Drack A, Sheffield VC. (1994) Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. *Arch. Ophthalmol.* 112:765-772.
3. Edwards AO, Miedziak A, Vrabec T, Verhoeven J, Acott TS, Weleber RG, Donoso LA. (1999) Autosomal dominant Stargardt-like macular dystrophy: I. Clinical characterization, longitudinal follow-up, and evidence for a common ancestry in families linked to chromosome 6q14. *Am. J. Ophthalmol.* 127:426-435.

4. Griesinger IB, Sieving PA, Ayyagari R. (2000) Autosomal dominant macular atrophy at 6q14 excludes CORD7 and MCDR1/PBCRA loci. *Invest. Ophthalmol. Vis. Sci.* 41:248-255.

5. Lagali PS, MacDonald IM, Griesinger IB, Chambers ML, Ayyagari R, Wong PW. (2000) Autosomal dominant Stargardt-like macular dystrophy segregating in a large Canadian family. *Can. J. Ophthalmol.* 35:315-24.

6. Toke DA, Martin CE. (1996) Isolation and characterization of a gene affecting fatty acid elongation in Saccharomyces cerevisiae. *J. Biol. Chem.* 271:18413-18422.

7. Oh CS, Toke DA, Mandala S, Martin CE. (1997) *ELO2* and *ELO3*, homologues of the saccharomyces cerevisiae *ELO1* gene, function in fatty acid elongation and are required for sphingolipid formation. *J. Biol. Chem.* 272:17376-17384.

8. Zhang K, Bither PP, Park R, Donoso LA, Seidman JG, Seidman CE. (1994) A dominant Stargardt's macular dystrophy locus maps to chromosome 13q34. *Arch. Ophthalmol.* 112:759-764.

9. Lagali PS, Kakuk LE, Griesinger IB, Wong PW, Ayyagari R. (2002) Identification and characterization of *C6orf37*, a novel candidate human retinal disease gene on chromosome 6q14. *Biochem Biophys Res Commun.* 293:356-365.

10. Thompson JD, Higgins DG, Gibson TJ. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22:4673-4680.

11. Kyte J, Doolittle RF. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.

12. Hofmann K, Stoffel W. (1993) TMbase - A database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler* 374:166.

13. Shah AB, Chernov I, Zhang HT, Ross BM, Das K, Lutsenko S, Parano E, Pavone L, Evgrafov O, Ivanova-Smolenskaya IA, Anneren G, Westermark K, Urrutia FH, Penchaszadeh GK, Sternlieb I, Scheinberg IH, Gilliam TC, Petrukhin K. (1997)

Identification and analysis of mutations in the Wilson disease gene (*ATP7B*): population frequencies, genotype-phenotype correlation, and functional analyses. *Am. J. Hum. Genet.* 61:317-328.

14. Tvrdik P, Asadi A, Kozak LP, Nedergaard J, Cannon B, Jacobsson A. (1997) *Cig30*, a mouse member of a novel membrane protein gene family, is involved in the recruitment of brown adipose tissue. *J. Biol. Chem.* 272:31738-31746.

15. Tvrdik P, Westerberg R, Silve S, Asadi A, Jakobsson A, Cannon B, Loison G, Jacobsson A. (2000) Role of a new mammalian gene family in the biosynthesis of very long chain fatty acids and sphingolipids. *J. Cell. Biol.* 149:707-718.

16. Shanklin J, Whittle E, Fox BG. (1994) Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* 33:12787-12794.

17. Fox BG, Shanklin J, Ai J, Loehr TM, Sanders-Loehr J. (1994) Resonance Raman evidence for an Fe-O-Fe center in stearoyl-ACP desaturase. Primary sequence identity with other diiron-oxo proteins. *Biochemistry* 33:12776-12786.

18. Jackson MR, Nilsson T, Peterson PA. (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO J.* 9:3153-3162.

19. Jackson MR, Nilsson T, Peterson PA. (1993) Retrieval of transmembrane proteins to the endoplasmic reticulum. *J. Cell. Biol.* 121:317-333.

20. Teasdale RD, Jackson MR. (1993) Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the golgi apparatus. *Annu. Rev. Cell. Dev. Biol.* 12:27-54.

21. Cinti DL, Cook L, Nagi MN, Suneja SK. (1992) The fatty acid chain elongation system of mammalian endoplasmic reticulum. *Prog. Lipid Res.* 31:1-51.

22. Allikmets R, Seddon JM, Bernstein PS, Hutchinson A, Atkinson A, Sharma S, Gerrard B, Li W, Metzker ML, Wadelius C, Caskey CT, Dean M, Petrukhin K. (1999) Evaluation of the Best disease gene in patients with age-related macular degeneration and other maculopathies. *Hum. Genet.* 104:449-453.

23. Millar AA, Kunst L. (1997) Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. *Plant J.* 12:121-131.

24. Millar AA, Clemens S, Zachgo S, Giblin EM, Taylor DC, Kunst L. (1999) *CUT1*, an *Arabidopsis* gene required for cuticular wax biosynthesis and pollen fertility, encodes a very-long-chain fatty acid condensing enzyme. *Plant Cell* 11:825-838.

25. Todd J, Post-Beittenmiller D, Jaworski JG. (1999) *KCS1* encodes a fatty acid elongase 3-ketoacyl-CoA synthase affecting wax biosynthesis in *Arabidopsis thaliana*. *Plant J.* 17:119-130.

26. Zank TK, Zahringer U, Lerchl J, Heinz E. (2000) Cloning and functional expression of the first plant fatty acid elongase specific for Δ^6-polyunsaturated fatty acids. *Biochem. Soc. Trans.* 28:654-658.

27. Das T, Thurmond JM, Bobik E, Leonard AE, Parker-Barnes JM, Huang YS, Mukerji P. (2000) Polyunsaturated fatty acid-specific elongation enzymes. *Biochem. Soc. Trans.* 28:658-660.

28. Beaudoin F, Michaelson LV, Hey SJ, Lewis MJ, Shewry PR, Sayanova O, Napier JA. (2000) Heterologous reconstitution in yeast of the polyunsaturated fatty acid biosynthetic pathway. *Proc. Natl. Acad. Sci. U. S. A.* 97:6421-6426.

29. Leonard AE, Bobik EG, Dorado J, Kroeger PE, Chuang LT, Thurmond JM, Parker-Barnes JM, Das T, Huang YS, Mukerji P. (2000) Cloning of a human cDNA encoding a novel enzyme involved in the elongation of long-chain polyunsaturated fatty acids. *Biochem. J.* 350:765-770.

30. Silve S, Leplatois P, Josse A, Dupuy PH, Lanau C, Kaghad M, Dhers C, Picard C, Rahier A, Taton M, Le Fur G, Caput D, Ferrara P, Loison G. (1996) The immunosuppressant SR 31747 blocks cell proliferation by inhibiting a steroid isomerase in Saccharomyces cerevisiae. *Mol. Cell. Biol.* 16:2719-2727.

31. Watts JL, Browse J. (2002) Genetic dissection of polyunsaturated fatty acid synthesis in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 99:5854-5859.

32. Kniazeva M, Sieber M, McCauley S, Zhang K, Watts JL, Han M. (2003) Suppression of the ELO-2 FA elongation activity results in alterations of the fatty acid composition and multiple physiological defects, including abnormal ultradian rhythms, in *Caenorhabditis elegans*. *Genetics* 163:159-169.

33. Westerberg R, Tvrdik P, Undén AB, Månsson JE, Norlén L, Jakobsson A, Holleran WH, Asadi A, Flodby P, Toftgård R, Capecchi MR, Jacobsson A. (2004) Role for Elovl3 and fatty acid chain length in development of hair and skin function. *J. Biol. Chem.* 279:5621-5629.

34. Poulos A. (1995) Very long chain fatty acids in higher animals--a review. *Lipids* 30:1-14.

35. Cook HW, Clarke JTR, Spence MW. (1983) Concerted stimulation and inhibition of desaturation, chain elongation, and esterification of essential fatty acids by cultured neuroblastoma cells. *J. Biol. Chem.* 258:7586-7591.

36. Fliesler SJ, Anderson RE. (1983) Chemistry and metabolism of lipids in the vertebrate retina. *Prog. Lipid. Res.* 22:79-131.

37. Sprecher H, Luthria DL, Mohammed BS, Baykousheva SP. (1995) Reevaluation of the pathways for the biosynthesis of polyunsaturated fatty acids. *J. Lipid Res.* 36:2471-2477.

38. Luthria DL, Mohammed BS, Sprecher H. (1996) Regulation of the biosynthesis of 4,7,10,13,16,19-docosahexaenoic acid. *J. Biol. Chem.* 271:16020-16025.

39. Moon YA, Shah NA, Mohapatra S, Warrington JA, Horton JD. (2001) Identification of a mammalian long chain fatty acyl elongase regulated by sterol regulatory element-binding proteins. J. Biol. Chem. 276:45358-45366.

40. Moon YA, Horton JD. (2003) Identification of two mammalian reductases involved in the two-carbon fatty acyl elongation cascade. *J. Biol. Chem.* 278:7335-7343.

41. Los DA, Murata N. (1998) Structure and expression of fatty acid desaturases. *Biochim. Biophys. Acta* 1394:3-15.

42. Brown MF. (1994) Modulation of rhodopsin function by properties of the membrane bilayer. *Chem. Phys. Lipids* 73:159-180.

43. Ahn J, Wong JT, Molday RS. (2000) The effect of lipid environment and retinoids on the ATPase activity of ABCR, the photoreceptor ABC transporter responsible for Stargardt macular dystrophy. *J. Biol. Chem.* 275:20399-20405.

44. Neuringer M. (2000) Infant vision and retinal function in studies of dietary long-chain polyunsaturated fatty acids: methods, results, and implications. *Am. J. Clin. Nutr.* 71:256S-267S.

45. Das UN. (2003) Long-chain polyunsaturated fatty acids in the growth and development of the brain and memory. *Nutrition* 19:62-65.

46. Harbige LS. (2003) Fatty acids, the immune response, and autoimmunity: a question of n-6 essentiality and the balance between n-6 and n-3. *Lipids* 38:323-341.

47. Herrera E. (2002) Implications of dietary fatty acids during pregnancy on placental, fetal and postnatal development--a review. *Placenta* 23:S9-S19.

48. Mandal MN, Ambasudhan R, Wong PW, Gage PJ, Sieving PA, Ayyagari R. (2004) Characterization of mouse orthologue of ELOVL4: genomic organization and spatial and temporal expression. *Genomics* 83:626-635.

49. Roepman R, Bernoud-Hubac N, Schick DE, Maugeri A, Berger W, Ropers HH, Cremers FP, Ferreira PA. (2000) The retinitis pigmentosa GTPase regulator (RPGR) interacts with novel transport-like proteins in the outer segments of rod photoreceptors. *Hum. Mol. Genet.* 9:2095-2105.

50. Daiger SP. RetNet, the Retinal Information Network http://www.sph.uth.tmc.edu/Retnet/ .

51. Donoso LA, Frost AT, Stone EM, Weleber RG, MacDonald IM, Hageman GS, Cibis GW, Ritter R 3rd, Edwards AO. (2001) Autosomal dominant Stargardt-like macular dystrophy: founder effect and reassessment of genetic heterogeneity. *Arch. Ophthalmol.* 119:564-570.

52. Vrabec TR, Tantri A, Edwards A, Frost A, Donoso LA. (2003) Autosomal dominant Stargardt-like macular dystrophy: identification of a new family with a mutation in the *ELOVL4* gene. *Am. J. Ophthalmol.* 136:542-545.

53. Wiedmann TS, Pates RD, Beach JM, Salmon A, Brown MF. (1988) Lipid-protein interactions mediate the photochemical function of rhodopsin. *Biochemistry* 27:6469-6474.

54. Chen Y, Houghton LA, Brenna JT, Noy N. (1996) Docosahexaenoic acid modulates the interactions of the interphotoreceptor retinoid-binding protein with 11-cis-retinal. *J. Biol. Chem.* 271:20507-20515.

55. Allikmets R, Singh N, Sun H, Shroyer NF, Hutchinson A, Chidambaram A, Gerrard B, Baird L, Stauffer D, Peiffer A, Rattner A, Smallwood P, Li Y, Anderson KL, Lewis RA, Nathans J, Leppert M, Dean M, Lupski JR. (1997) A photoreceptor cell-specific ATP-binding transporter gene (*ABCR*) is mutated in recessive Stargardt macular dystrophy. *Nature Genet.* 15:236-246.

56. Weng J, Mata NL, Azarian SM, Tzekov RT, Birch DG, Travis GH. (1999) Insights into the function of Rim protein in photoreceptors and etiology of Stargardt's disease from the phenotype in *abcr* knockout mice. *Cell* 98:13-23.

57. Sun, H, Molday RS, Nathans J. (1999) Retinal stimulates ATP hydrolysis by purified and reconstituted ABCR, the photoreceptor-specific ATP-binding cassette transporter responsible for Stargardt disease. *J. Biol. Chem.* 274:8269-8281.

58. Gelehrter TD, Collins FS, Ginsburg D. (1998) *Principles of Medical Genetics.* Baltimore, MD:Williams and Wilkins. p.143.

59. Wiegand RD, Giusto NM, Rapp LM, Anderson RE. (1983) Evidence for rod outer segment lipid peroxidation following constant illumination of the rat retina. *Invest. Ophthalmol. Vis. Sci.* 24:1433-1435.

60. Beatty S, Koh H, Phil M, Henson D, Boulton M. (2001) The role of oxidative stress in the pathogenesis of age-related macular degeneration. *Surv. Ophthalmol.* 45:115-134.

61. Bernstein PS, Tammur J, Singh N, Hutchinson A, Dixon M, Pappas CM, Zabriskie NA, Zhang K, Petrukhin K, Leppert M, Allikmets R. (2001) Diverse macular dystrophy phenotype caused by a novel complex mutation in the *ELOVL4* gene. *Invest. Ophthalmol. Vis. Sci.* 42:3331-3336.

62. Zhang K, Kniazeva M, Hutchinson A, Han M, Dean M, Allikmets R. (1999) The *ABCR* gene in recessive and dominant Stargardt diseases: a genetic pathway in macular degeneration. *Genomics* 60:234-237.

63. Kajiwara K, Berson EL, Dryja TP. (1994) Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/*RDS* and *ROM1* loci. *Science* 264:1604-1608.

64. van Kuijk FJ, Buck P. (1992) Fatty acid composition of the human macula and peripheral retina. *Invest. Ophthalmol. Vis. Sci.* 33:3493-3496.

65. De La Paz M, Anderson RE. (1992) Region and age-dependent variation in susceptibility of the human retina to lipid peroxidation. *Invest. Ophthalmol. Vis. Sci.* 33:3497-3499.

66. Snodderly DM, Sandstrom MM, Leung IY, Zucker CL, Neuringer M. (2002) Retinal pigment epithelial cell distribution in central retina of rhesus monkeys. *Invest. Ophthalmol. Vis. Sci.* 43:2815-2818.

67. Pacione LR, Szego MJ, Ikeda S, Nishina PM, McInnes RR. (2003) Progress toward understanding the genetic and biochemical mechanisms of inherited photoreceptor degenerations. *Annu. Rev. Neurosci.* 26:657-700.

Figure 5.1 Strategy used to identify and define the human *ELOVL4* gene. Refer to text for details. *A*, Physical map of the STGD3/adMD genomic interval depicting the relative positions of the human chromosome 6q microsatellite markers in the critical region. *cen*, centromere; *tel*, telomere. *B*, The Sanger Center contig Chr_6ctg16 comprised of overlapping PAC clones in the disease gene region. *C*, Position of candidate genes in the critical region that were tested for disease-associated mutations in STGD3/adMD families. *D*, Positions of two human ESTs mapping to the critical region, displaying retinal and RPE expression by macroarray analysis, and corresponding to the *ELOVL4* gene transcript. GenBank accession numbers for each of the EST clones are given. *E*, Genomic organization of the human *ELOVL4* gene. Rectangles represent exons and intervening horizontal lines represent introns. Open rectangles correspond to the open reading frame, while shading denotes untranslated regions (UTR). The lengths of exon sequences are given within the rectangles, and approximate sizes of the introns are indicated below the intervening horizontal lines. The positions of the start (ATG) and stop (TAA) codons of the ORF are shown. Asterisks represent consensus polyadenylation signals (AATAAA) found at positions 2123 and 2964 of the ELOVL4 mRNA sequence. Contributions to data for Panels A-C were made by Dr. Radha Ayyagari (University of Michigan), Dr. Kang Zhang (Johns Hopkins University), Dr. Konstantin Petrukhin (Merck Research Laboratories), and associates.

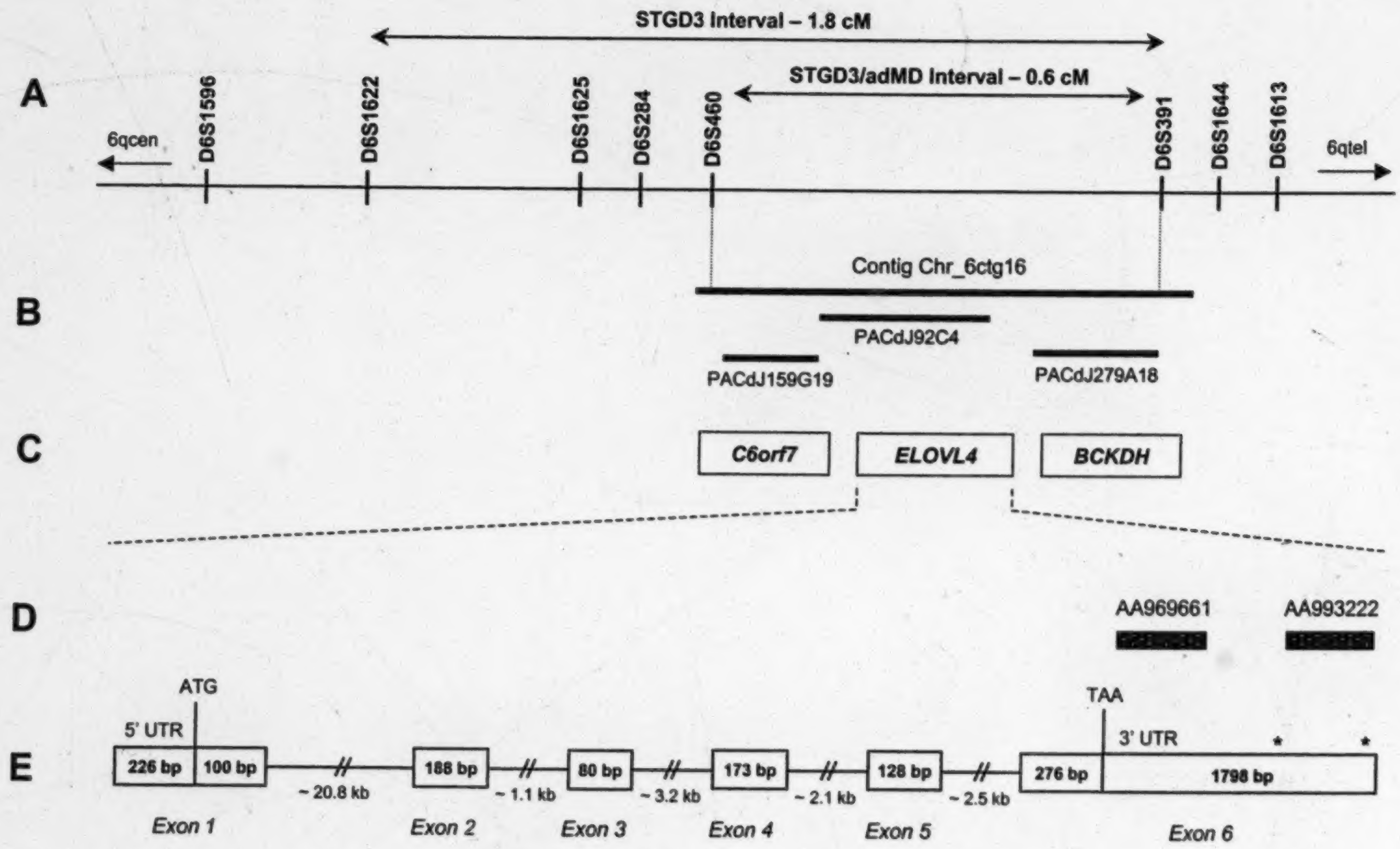
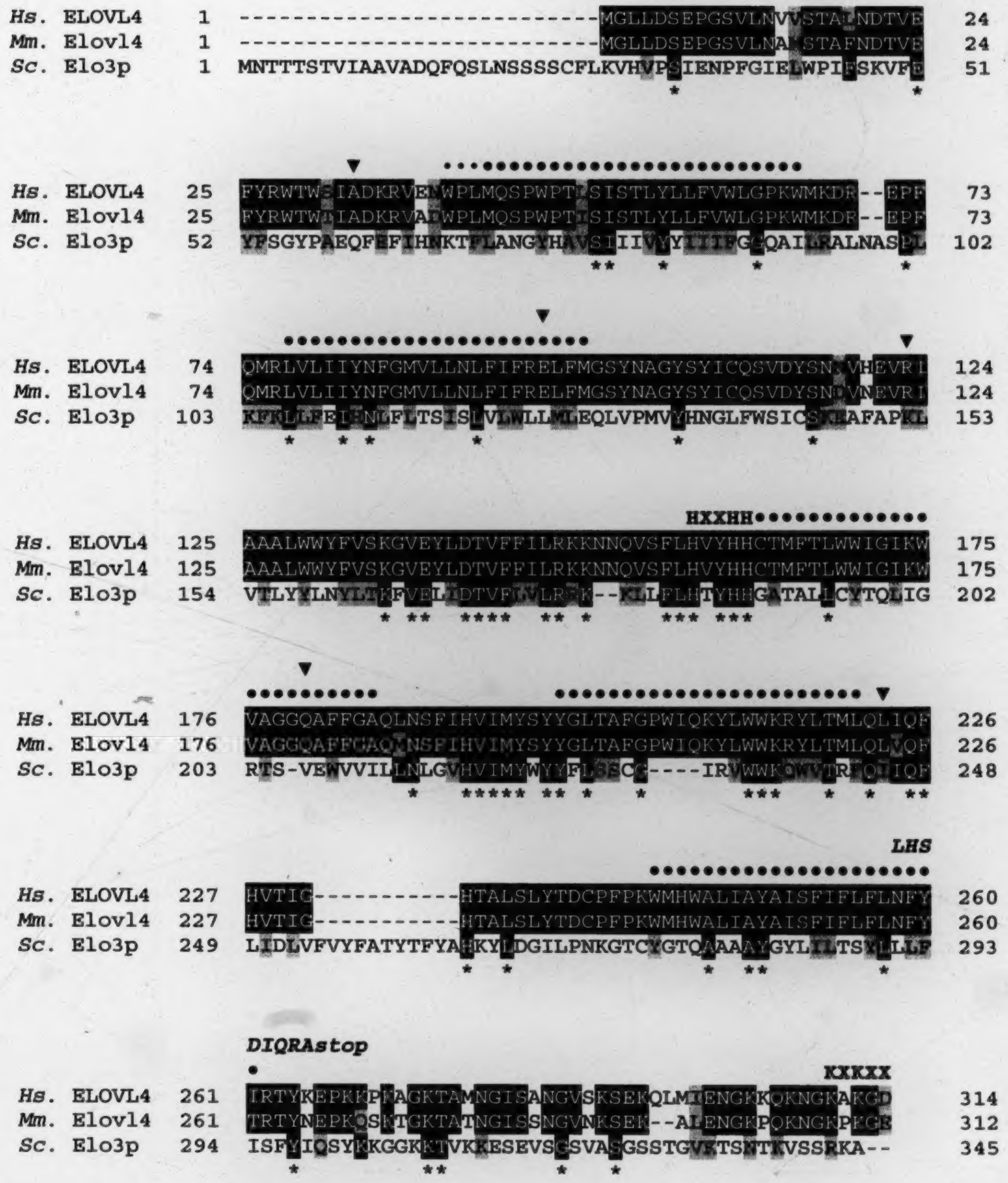

A
STGD3 Interval – 1.8 cM
STGD3/adMD Interval – 0.6 cM
6qcen
D6S1596
D6S1622
D6S1625
D6S284
D6S460
D6S391
D6S1644
D6S1613
6qtel
B
Contig Chr_6ctg16
PACdJ92C4
PACdJ159G19
PACdJ279A18
C
C6orf7
ELOVL4
BCKDH
D
AA969661
AA993222
E
ATG
5' UTR
226 bp
100 bp
~ 20.8 kb
188 bp
~ 1.1 kb
80 bp
~ 3.2 kb
173 bp
~ 2.1 kb
128 bp
~ 2.5 kb
TAA
3' UTR
276 bp
1798 bp
*
*
Exon 1
Exon 2
Exon 3
Exon 4
Exon 5
Exon 6

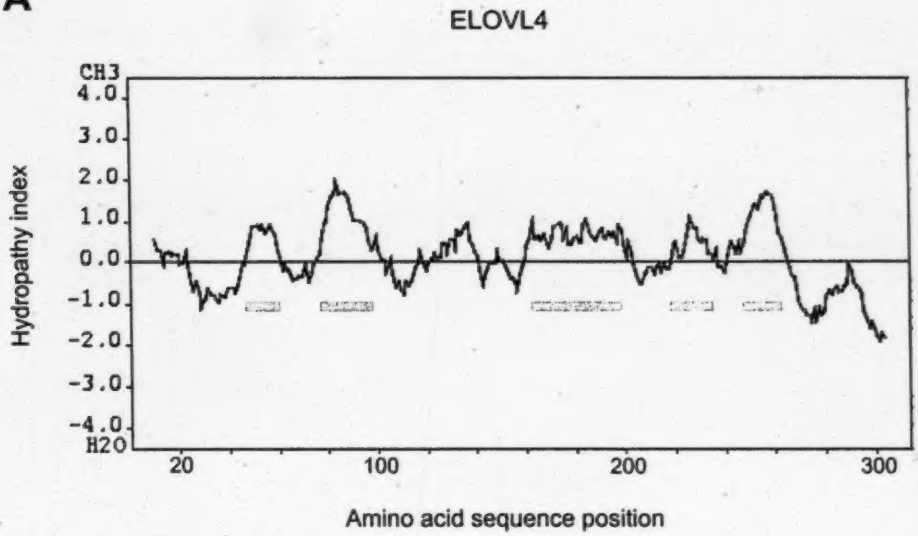
Figure 5.2 Comparison of human ELOVL4 and mouse Elovl4 protein sequences with the product of the yeast *ELO3* gene, Elo3p. The amino acid sequences of the *Homo sapiens* (*Hs.*; GenBank accession number AAH38506), *Mus musculus* (*Mm.*; GenBank accession number AAH37030), and *Saccharomyces cerevisiae* (*Sc.*; GenBank accession number P40319) homologues were aligned using the Clustal W program (version 1.83). Hyphens represent gaps introduced to maintain the alignment. The amino acid residues are numbered to the left and right of the sequences. Identical residues among the sequences in the alignment are shown on a black background, and conservative amino acid substitutions are shaded in grey. Residues identical among all three sequences are marked with asterisks. Vertical arrowheads indicate boundaries between exons 1-6 in the human *ELOVL4* gene. Predicted positions of the five putative transmembrane domains are indicated by bullets. The carboxy-terminal dilysine signal (KXKXX) responsible for intracellular targeting to the endoplasmic reticulum and the diiron-oxo-binding ‘histidine box’ (HXXHH) motif are indicated in bold above the protein sequences. The truncated carboxyl terminus caused by the 5-bp deletion in *ELOVL4* is shown in bold and italics above the human protein sequence.



```
Hs. ELOVL4    1   ----------------------------MGLLDSEPGSVLNVVSTALNDTVE   24
Mm. Elovl4    1   ----------------------------MGLLDSEPGSVLNAMSTAFNDTVE   24
Sc. Elo3p     1   MNTTTSTVIAAVADQFQSLNSSSSCFLKVHVPSIENPFGIELWPIFSKVFE   51

Hs. ELOVL4   25   FYRWTWSIADKRVENWPLMQSPWPTLSISTLYLLFVWLGPKWMKDR--EPF   73
Mm. Elovl4   25   FYRWTWTIADKRVADWPLMQSPWPTISISTLYLLFVWLGPKWMKDR--EPF   73
Sc. Elo3p    52   YFSGYPAEQFEFIHNKTFLANGYHAVSIIIVYYIIIFGGQAILRALNASPL  102

Hs. ELOVL4   74   QMRLVLIIYNFGMVLLNLFIFRELFMGSYNAGYSYICQSVDYSNNVHEVRI  124
Mm. Elovl4   74   QMRLVLIIYNFGMVLLNLFIFRELFMGSYNAGYSYICQSVDYSNDVNEVRI  124
Sc. Elo3p   103   KFKLLFEIHNLFLTSISLVLWLLMLEQLVPMVYHNGLFWSICSKEAFAPKL  153

                                                  HXXHH
Hs. ELOVL4  125   AAALWWYFVSKGVEYLDTVFFILRKKNNQVSFLHVYHHCTMFTLWWIGIKW  175
Mm. Elovl4  125   AAALWWYFVSKGVEYLDTVFFILRKKNNQVSFLHVYHHCTMFTLWWIGIKW  175
Sc. Elo3p   154   VTLYYLNYLTKFVELIDTVFLVLRRK--KLLFLHTYHHGATALLCYTQLIG  202

Hs. ELOVL4  176   VAGGQAFFGAQLNSFIHVIMYSYYGLTAFGPWIQKYLWWKRYLTMLQLIQF  226
Mm. Elovl4  176   VAGGQAFFGAQLNSFIHVIMYSYYGLTAFGPWIQKYLWWKRYLTMLQLVQF  226
Sc. Elo3p   203   RTS-VEWVVILLNLGVHVIMYWYYFLSSCG----IRVWWKQWVTRFQIIQF  248

                                                                LHS
Hs. ELOVL4  227   HVTIG------------HTALSLYTDCPFPKWMHWALIAYAISFIFLFLNFY  260
Mm. Elovl4  227   HVTIG------------HTALSLYTDCPFPKWMHWALIAYAISFIFLFLNFY  260
Sc. Elo3p   249   LIDLVFVYFATYTFYAHKYLDGILPNKGTCYGTQAAAAYGYLILTSYLLLF  293

                  DIQRAstop
                                                             KXKXX
Hs. ELOVL4  261   IRTYKEPKKPKAGKTAMNGISANGVSKSEKQLMIENGKKQKNGKAKGD    314
Mm. Elovl4  261   TRTYNEPKQSKTGKTATNGISSNGVNKSEK--ALENGKPQKNGKPKGE    312
Sc. Elo3p   294   ISFYIQSYKKGGKKTVKKESEVSGSVASGSSTGVKTSNTKVSSRKA--    345
```

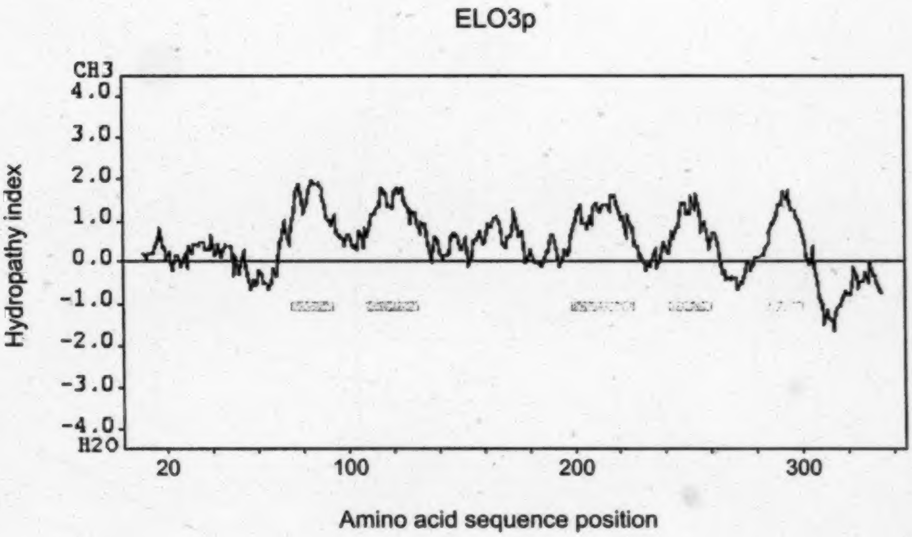
Figure 5.3 Hydropathy profile analysis of the human ELOVL4 and yeast Elo3p protein sequences. *A*, Hydropathy plot for *H. sapiens* ELOVL4 (GenBank accession number AAH38506); *B*, Hydropathy plot for *S. cerevisiae* Elo3p (GenBank accession number P40319). Regions above the horizontal bars represent amino acids with a positive hydropathy index. Five predicted membrane-spanning regions, based on the algorithm of Kyte and Doolittle,[11] are underlined in grey. Note the striking similarity in hydropathy plots between the ELOVL4 and Elo3p proteins despite a relatively low level of overall sequence identity (refer to Figure 5.2).

A



ELOVL4

B



ELO3p

Figure 5.4 Expression profile of the *ELOVL4* gene in human tissues. Northern blots of human RNA samples probed with ELOVL4 and control (β-actin) cDNA probes are shown. Choroid, ciliary body, retina, and retinal pigment epithelium (RPE) lanes contain 5 μg of total RNA while all other lanes contain 2 μg of poly(A)$^+$ RNA. Positions of RNA size markers are indicated on the right (*kb*, kilobases).

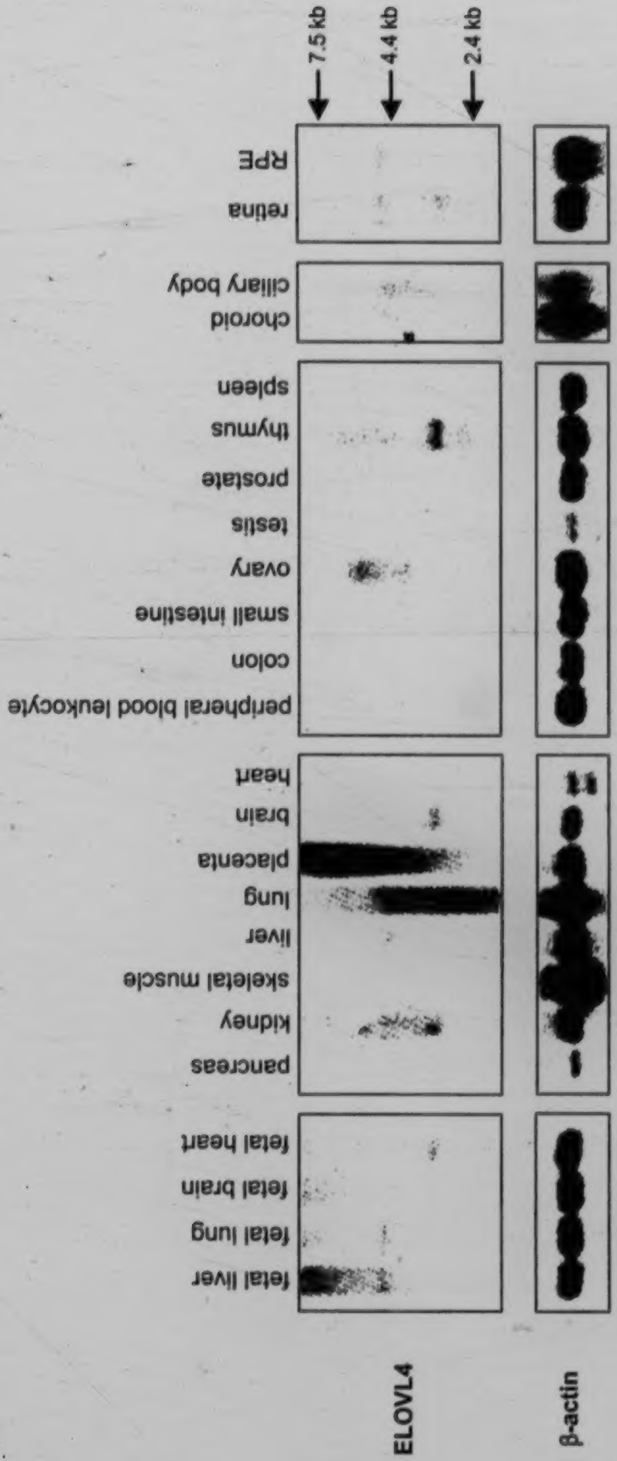
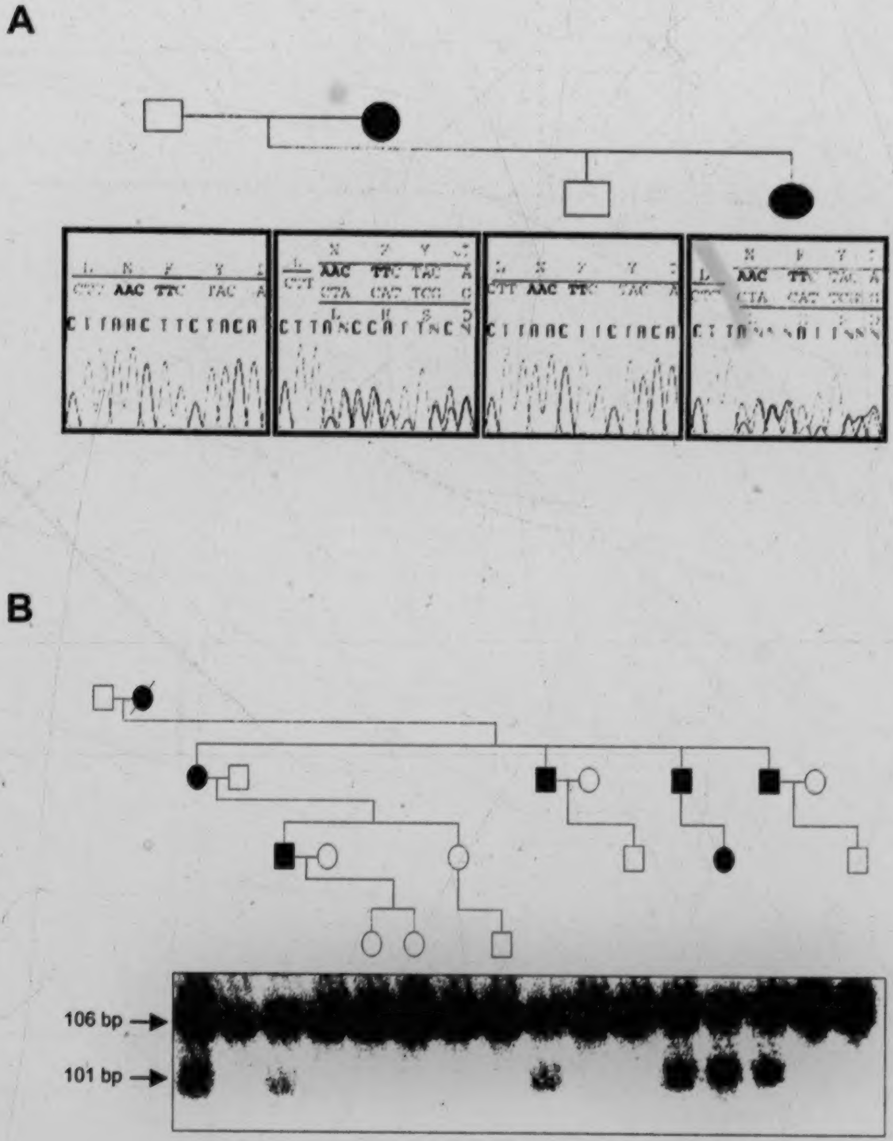

ELOVL4
β-actin
fetal liver
fetal lung
fetal brain
fetal heart
pancreas
kidney
skeletal muscle
liver
lung
placenta
brain
heart
peripheral blood leukocyte
colon
small intestine
ovary
testis
prostate
thymus
spleen
choroid
ciliary body
retina
RPE
7.5 kb
4.4 kb
2.4 kb

Table 5.1 Primer pairs used for PCR amplification of *ELOVL4* exons for mutation analysis in STGD3/adMD families. *Primers used for detection of the 5-bp deletion in a smaller fragment of exon 6 by polyacrylamide gel electrophoresis; the amplification product of the wild-type allele is 106-bp in length while that for the mutant allele is 101-bp in length.

Exon	Forward Primer	Reverse Primer	Expected PCR Product Size (bp)
1	5'-TCTCTCTGGGTCTCCGCTT-3'	5'- GAGGGGAGGCCTTAACATTC-3'	346
2	5'-AGCCACTTGCAGGAGTCAGT-3'	5'-TGATGGTTTTACACATTCTCATTTTT-3'	451
3	5'-AGCAATCGGAATGCATGAAA-3'	5'-TTTTCACAGACTGGGGCCTA-3'	304
4	5'-TCCATGCCTTGTACATTTTGTG-3'	5'-GGTAGATCAAGTCAAAGTCC-3'	680
5	5'-ACATCTCAGTGGCTTACTGCCT-3'	5'-GAAAATTGTCTAAAATGACATTGCAC-3'	290
6	5'-GAAGATGCCGATGTTGTTAAAAG-3'	5'-GTCAACAACAGTTAAGGCCCA-3'	376
6*	5'-GCAATATTCCCCAAATGGATGCACTGG-3'	5'-TTGCGGCCGCTTAGGCTCTTTGTATGTCCGAAT-3'	106 or 101

Figure 5.5 Segregation of the 790-794delAACTT mutation in families with the STGD3 and adMD phenotypes. *A*, Sequence of exon 6 in a small branch of family B. Affected mother and daughter are heterozygous for mutant and wild-type alleles, whereas the unaffected father and son are homozygous for the wild-type allele. The nucleotide sequences and translations are shown above the direct sequencing traces for exon 6 PCR products. The 5-base pair deletion detected in affected individuals is shown in bold. *B*, Co-segregation of the 790-794delAACTT mutation in a branch of the Canadian STGD3 family. Exon 6 fragments amplified by PCR were resolved by electrophoresis in a 6% polyacrylamide gel. The affected individuals who are heterozygous for the 5-bp deletion have two PCR bands corresponding to the wild-type allele (106-bp band) and the mutant allele (101-bp band), whereas normal individuals have only one wild-type band of 106-bp. Panel A image provided by Dr. Kang Zhang.


A
CTT AAC TTC TAC A
CTTAACTTCTACA
B
106 bp
101 bp

Chapter 6

Evolutionarily Conserved *ELOVL4* Gene Expression in the Vertebrate Retina

A version of this chapter has been published.
Lagali PS, Liu J, Ambasudhan R, Kakuk LE, Bernstein SL, Seigel GM, Wong PW, Ayyagari R. Evolutionarily Conserved *ELOVL4* Gene Expression in the Vertebrate Retina. *Investigative Ophthalmology and Visual Science* 2003;44(7):2841-2850.

6.1 INTRODUCTION

Stargardt macular dystrophies are inherited disorders of the retina that are characterized by impaired central vision, cellular atrophy of retinal and retinal pigment epithelial cells initiating in the macular region, and the presence of yellow fundus flecks surrounding the atrophic area resulting from the sub-retinal accumulation of lipofuscin-like material.[1-4] Both autosomal recessive and autosomal dominant forms of Stargardt disease (STGD) have been described.[5,6]

Classical or recessive Stargardt disease (STGD1; OMIM #248200; http://www.ncbi.nlm.nih.gov/omim) is the most common form of hereditary macular dystrophy, occurring at an incidence of approximately 1 in 10,000 live births, and accounts for roughly 7% of all retinal dystrophies.[7] A distinguishing clinical feature of the disease is the dark or silent choroid observed upon fluorescein angiography of the fundus resulting from the accumulation of lipofuscin. Mutations in the *ABCA4* (*ABCR*) gene on human chromosome 1p32 have been found to underlie STGD1.[8,9] This gene encodes a retina-specific adenosine triphosphate-binding cassette transport protein found in photoreceptor cell outer segment disc membranes and functions as an outwardly directed flippase to facilitate the shuttling of N-retinylidene-phosphatidylethanolamine from the interior to the exterior of the outer segments for subsequent retinoid recycling by RPE cells following phototransduction.[10,11] *ABCA4* mutations have also been implicated in the etiology of a variety of other retinal diseases including retinitis pigmentosa, cone-rod dystrophy, and also cases of age-related macular degeneration (AMD).[9,12,13] Cases of autosomal dominant Stargardt macular dystrophy (adSTGD) are comparatively rare and a number of genetic loci for diseases with typical adSTGD traits have been mapped, including STGD3 and adMD (OMIM #600110) on human chromosome 6q14[14,15] and STGD4 (OMIM #603786) on chromosome 4p.[16] Autosomal dominant forms of Stargardt macular dystrophy are clinically quite similar to the phenotype characteristic of recessive Stargardt disease, however there exist a number of key differentiating features. The most notable differences between STGD1 and STGD3/adMD-associated phenotypes in particular are the later and more variable age of onset and the absence of the dark choroid observed in fluorescein angiograms of dominant Stargardt patients.[15,17,18]

Through a positional candidate approach, the gene defect corresponding to the STGD3/adMD locus responsible for the disease phenotype in five related pedigrees throughout North America was recently identified.[19] The mutation, a 5-base pair deletion, lies in the *ELOVL4* gene, a novel member of a family of genes that encode enzymatic components of the microsomal fatty acid elongation system referred to as fatty acid elongases (ELOs). The *ELOVL4* gene consists of 6 exons and encodes a 2.9-kilobase full-length mRNA transcript, which in turn codes for a 314-amino acid protein that contains the three typical features that define a fatty acid elongase[20,21]: five putative transmembrane segments, a single HXXHH 'histidine box' redox center motif, and a carboxy-terminal dilysine targeting signal thought to be responsible for the retrieval of transmembrane proteins to the endoplasmic reticulum (ER) from the Golgi apparatus.[22] The 5-bp deletion in the gene identified in dominant Stargardt/adMD-affected pedigrees causes a reading frame shift that results in premature termination of the translation product, leading to substitutions in amino acids 264-271 and the loss of the 51 carboxy-terminal residues in the mutant protein. This mutation has subsequently been verified in affected members present in other branches of the same family.[23,24] In addition, a novel complex mutation in the *ELOVL4* gene has been described in an unrelated family that presents with diverse macular dystrophy phenotypes.[25]

Thus it is clearly established that the *ELOVL4* gene plays a critical role in the pathogenesis of macular degeneration. However the precise nature of this role is at present unknown. Clues to the importance of this gene in retinal cell biology have also come from the identification of both mouse and yeast homologues, implying evolutionarily conserved gene expression. This limited assessment of conservation, its implied role in fatty acid metabolism, and its association with retinal degeneration suggests that *ELOVL4* is an essential retina-expressed gene. To further characterize this gene in normal retinal biology, this study describes the use of nucleic acid hybridization, bioinformatic, and immunodetection analyses to examine the sequence conservation and expression of ELOVL4 gene products across a wide range of species. The results demonstrate that *ELOVL4* represents a highly evolutionarily conserved gene that is expressed in the photoreceptor cell layer within the retina of every mammalian species examined.

6.2 MATERIALS AND METHODS

6.2.1 Animal Tissues and Donor Eyes

Human ocular tissue was obtained from eyes donated to the Michigan Eye Bank (Ann Arbor, MI) and through the National Disease Research Interchange (NDRI, Philadelphia, PA). Archival human retinal sections were obtained from the Ocular Pathology Lab at the University of Rochester. Rhesus monkey eyes were kindly provided by J. Cogan (Bureau of Biologics, Bethesda, MD). Ocular dissections were performed as previously described.[26] Mouse, Sprague-Dawley rat, and cat eyes were obtained from normal control animals in the University of Michigan animal facility and dissected immediately after harvesting. Bovine eyes were obtained from a local abattoir, stored on ice upon harvesting, and dissections were performed within 3 hours post-mortem. The care and handling of animals was in compliance with the Association of Research in Vision and Ophthalmology (ARVO) Statement for the Use of Animals in Ophthalmic and Vision Research.

6.2.2 Southern Blot Analysis

A commercially available zooblot containing 8 micrograms of *Eco*RI-digested genomic DNA isolated from various species (human, monkey, rat, mouse, dog, cow, rabbit, chicken, yeast) was purchased (BD Biosciences Clontech). A 376-base pair DNA probe corresponding to exon 6 of the human *ELOVL4* gene was generated from a genomic DNA template by the polymerase chain reaction using forward primer 5'-GAAGATGCCGATGTTGTTAAAAG-3' and reverse primer 5'-GTCAACAACAGTTAAGGCCCA-3'. PCR cycling parameters used were as follows: 1 cycle of denaturation at 94°C for 3 minutes; 30 cycles of denaturation at 94°C for 30 seconds, primer annealing at 55°C for 30 seconds, primer extension at 72°C for 40 seconds; 1 cycle of extension at 72°C for 3 minutes. A probe corresponding to *ELOVL4* coding regions was also generated by total cDNA synthesis using a commercial kit (Invitrogen Canada Inc.) and five micrograms of human retina total RNA as template, followed by PCR using 100 nanograms of total cDNA template with forward primer 5'-GTGTGGAAAATTGGCCTCTG-3' and reverse primer 5'-CATGGCTGTTTTTCCAGCTT-3' to amplify a 743-base pair cDNA fragment containing protein-encoding sequences found within exons 2 through 6. PCR cycling parameters used were as follows: 1 cycle of denaturation at 94°C for 3 minutes; 30 cycles of denaturation at 94°C for 1 minute, primer annealing at 45°C for 1 minute, primer extension at 72°C for 3 minutes; 1 cycle of extension at 72°C for 7 minutes. Probes were purified using the QIAquick Gel Extraction Kit (QIAGEN Inc.) and radiolabeled with [α^{32}P]-dCTP (Amersham Biosciences) by random primer labeling using the Multiprime DNA Labeling System (Amersham Biosciences). Hybridization of each probe to the zooblot was carried out in Hybrisol II hybridization solution (Serologicals Corporation) at 42°C for 16 hours using 10^6 counts per minute (cpm) of labeled, purified probes per milliliter of hybridization solution. The blot was then washed twice for 15 minutes in 2X SSC, once for 30 minutes in 2X SSC, 0.1% SDS, and once for 10 minutes in 0.1% SSC, 0.1% SDS all at 42°C, followed by exposure to Kodak X-Omat film at -70°C for 24 hours.

6.2.3 RNA Isolation and Northern Blot Analysis

Total RNA was isolated from primate ocular tissues and retinal tissue derived from various mammalian species using the Trizol reagent (Invitrogen Canada Inc.) according to the manufacturer's protocol. Five micrograms of each total RNA sample was electrophoresed on a 1% agarose-formaldehyde gel and subsequently blotted onto a nylon membrane (Pall Corporation) via capillary transfer. Northern blots were consecutively hybridized with probes corresponding to human ELOVL4 and β-actin coding regions. The ELOVL4 probe used was the 743-base pair cDNA fragment described above. The β-actin probe was amplified by PCR from a plasmid clone with forward primer 5'-GTCCACACCCGCCGCCAGCTCACCAT-3' and reverse primer 5'-ACTGGTCTCAAGTCAGTGTACAGGTA-3' to generate a 1698-base pair nearly full-length cDNA product. Radiolabeled probes were prepared and hybridizations and washes were carried out as described above at either 42°C or 55°C for blots containing multi-species RNA samples, or at 65°C for primate RNA blots. For increased specificity of probe-binding, multi-species blots were blocked with 0.5 micrograms of denatured human Cot-1 DNA (Invitrogen Canada Inc.) added to the hybridization solution. ELOVL4-probed blots were exposed to film for 18-48 hours.

6.2.4 *Production of ELOVL4 Antibodies*

Anti-ELOVL4 antibodies were developed by Dr. Radha Ayyagari (University of Michigan, Ann Arbor, MI). A synthetic 15-amino acid peptide (SIADKRVENWPLMQS) corresponding to amino acids 31-45 of the human ELOVL4 protein was conjugated to keyhole limpet hemocyanin and subcutaneously injected into rabbits for the production of polyclonal antibodies (Research Genetics, Inc.). Pre-immune serum was also collected from the rabbits prior to immunization for use in control experiments. Antibodies were affinity-purified using a column containing immobilized peptide antigen.

6.2.5 *Preparation of Retinal Protein Extracts*

Volumes of retinal tissue samples were estimated and an equal volume of 2X lysis buffer (100 mM Tris-HCl pH 7.4, 300 mM NaCl, 1% Triton X-100, 1 Complete Protease Inhibitor Cocktail Tablet (Roche Diagnostics Corp., Indianapolis, IN) dissolved in 10 millilitres of water) was added. Tissues were homogenized using a Sonifier 250 sonicator (Branson Ultrasonics, Danbury, CT), with samples pulsed for 10 seconds at 2-minute intervals over 20 minutes at 4°C. Homogenates were spun at 5000 rpm for 3-5 minutes in a microcentrifuge to precipitate insoluble material, and the supernatants were further analyzed for protein content. Total protein concentration was estimated with the BCA Protein Assay Kit (Pierce Biotechnology, Rockford, IL). Protein extracts were prepared in part by Ms. Laura E. Kakuk (Univeristy of Michigan).

6.2.6 *SDS-PAGE and Immunoblot Analysis*

Retinal protein extracts (approximately 200 micrograms of total protein) were added to an equal volume of 2X SDS-PAGE loading buffer (100 mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol) and electrophoresed on a 12% polyacrylamide gel. Proteins were transferred to a nitrocellulose membrane (Sigma-Aldrich Corp., St. Louis, MO) by electroblotting using the Mini Trans-Blot Protein Blotting Cell (Bio-Rad Laboratories, Hercules, CA). Immunoblots were blocked with 3% skim milk powder for 1 hour at room temperature, followed by incubation with the ELOVL4 polyclonal antiserum diluted 1:1000 in 3% skim milk at room temperature for 1 hour in the presence or absence of 20 micrograms of competing ELOVL4 peptide. Blots were subsequently labeled with a 1:5000 dilution of horseradish peroxidase-conjugated goat anti-rabbit secondary antibody (Pierce Biotechnology) and subjected to immunodetection by enhanced chemiluminescence (ECL) using ECL Western Blotting Detection Reagents (Amersham Biosciences, Piscataway, NJ). SDS-PAGE and immunoblot analysis of retinal protein extracts were performed in part by Ms. Laura E. Kakuk.

6.2.7 Bioinformatic Analysis

ELOVL4 homologues and orthologues were identified by performing protein homology searches of sequence databases with the translated ELOVL4 mRNA sequence (GenBank accession number AF277094) and the BLASTP program at the NCBI website, http://www.ncbi.nlm.nih.gov/blastp.[27] Proteins identified by the BLASTP search program with greater than 25% identity to the human ELOVL4 amino acid sequence were classified as homologues. Sequence alignments and phylogenetic analysis were performed using the Clustal W (versions 1.8, 1.82)[28] and PHYLIP (*Phylogeny Inference Package*) (version 3.5)[29,30] programs and Command Line applications for sequence analysis and phylogeny available at the Canadian Bioinformatics Resource website, http://www.cbr-rbc.nrc-cnrc.gc.ca/.

6.2.8 Immunohistochemistry

6.2.8.1 Human Retina Sections

Whole human eyes were paraffin-embedded and cut into 4-micron tissue sections. Human tissue sections were permeabilized in 0.25% Triton X-100 for 5 minutes and rinsed in phosphate-buffered saline (PBS). Endogenous peroxidase was quenched with a 5-minute incubation in 3% hydrogen peroxide. Tissues were rinsed twice for 5 minutes in PBS and blocked for 15 minutes in 5% goat serum. Following two additional 5-minute rinses in PBS, sections were incubated with a 1:100 dilution of ELOVL4 polyclonal antiserum or pre-immune serum as a negative control for one hour at room temperature. Tissues were subsequently rinsed three times for 5 minutes in PBS and then incubated with 1 μg/ml of biotinylated goat anti-rabbit immunoglobulin secondary antibody (Vector Laboratories, Inc., Burlingame, CA) for 1 hour at room temperature. After a final PBS rinse, tissues were incubated with HRP-conjugated avidin (Vectastain Elite kit; Vector Laboratories, Inc.) for 20 minutes. Tissue sections were equilibrated in

0.05 M Tris-HCl pH 7.4, then developed with a diaminobenzidine (DAB) kit (Pierce Chemicals, Rockford, IL), and the brown- or black-coloured reaction product was visualized by light microscopy. For DNA counterstaining, sections were incubated with 30% methyl green for 50 seconds. Immunohistochemical analysis of human retinal tissue sections was performed by Dr. Gail M. Seigel (State University of New York at Buffalo, Buffalo, NY).

6.2.8.2 Animal Retina Sections

Mouse, cat, and rat specimens were perfused with 4% paraformaldehyde and eyes were extracted and fixed overnight at 4°C. Bovine eyes were obtained from an abattoir and immersed in 4% paraformaldehyde overnight. Anterior segment of the eye was removed. Eyecup was buffer rinsed and embedded in paraffin. Specimens were sectioned at 5 to 7 microns. Tissue sections were dewaxed in xylene and rinsed in PBS. Rehydrated sections were treated with proteinase K (500 ng/µl) as previously described.[31] Immunohistochemical analysis was performed using the VECTASTAIN ABC kit (Vector Laboratories, Inc.) following the manufacturer's protocol. Rat and cat sections were analyzed using the ELOVL4 polyclonal antiserum at a 1:400 dilution, and mouse and cow sections were analyzed with the ELOVL4 affinity-purified antibodies at a 1:10,000 dilution. Controls were performed using a 1:400 dilution of the pre-immune serum in place of the ELOVL4 antibodies. Results were visualized by light microscopy. Animal retina tissue sections were analyzed by Dr. Jiafan Liu and Dr. Rajesh Ambasudhan (University of Michigan).

6.3 RESULTS

6.3.1 Multi-Species Conservation of ELOVL4 DNA Sequences and mRNA Expression

To detect the presence of *ELOVL4* sequence homologues and/or orthologues in a number of different species, a Southern blot containing genomic DNA isolated from human, monkey, rat, mouse, dog, cow, rabbit, chicken, and yeast was probed with *ELOVL4*-derived sequences. Hybridization signals were detected for all the species examined (Figure 6.1), suggesting that the *ELOVL4* gene is conserved across a variety of eukaryotic species.

Retinal expression of the *ELOVL4* gene has already been demonstrated in human tissue samples.[19,23] To investigate the expression of ELOVL4 mRNA in additional human eye tissues, northern blot analysis was performed. The approximately 2.9-kilobase (kb) and 2.4-kb transcript sizes previously reported for the human ELOVL4 gene were consistently observed in different peripheral retina samples (Figure 6.2), however two additional transcripts of roughly 4.5-kb and 0.25-kb in size were also detected. These bands persisted even following Cot-1 DNA suppression and high stringency washing of blots. Similar hybridization signals were observed for human fovea tissue, however the 2.9-kb and 2.4-kb transcripts appear to be absent in RPE, choroid, and ciliary body tissue samples (Figure 6.2). This suggests that *ELOVL4* or *ELOVL4*-like genes are differentially expressed among different human eye tissues. All four transcripts were detected in rhesus monkey peripheral retina and macular tissues, indicating that *ELOVL4* gene expression is highly conserved between primate retinal tissues.

To extend these findings to additional species, retinal RNA samples derived from mouse, cow, and pig were similarly assessed by northern analysis (Figure 6.2). The approximately 4.5-kb transcript appears in all the species, while the 2.9-kb transcript is clearly observed in the mouse and cow retina but is undetectable in the pig retina RNA. The 2.4-kb transcript readily detected in the primate retina samples is not as easily observed in the retinal RNA from the other species under the stringent hybridization and washing conditions used. The approximately 0.25-kb transcript also appears in all samples except for pig retina. These results suggest that different species express a different complement of *ELOVL4* transcripts that may share varying degrees of sequence conservation, which in turn would lead to differences in detection capabilities based on the hybridization conditions used.

6.3.2 ELOVL4 Protein Orthologues and Homologues

The translated human ELOVL4 mRNA sequence (GenBank accession number AF277094) was used to search protein sequence databases to identify related proteins found in other species. Putative proteins exhibiting 25% or greater identity at the amino acid level with the human ELOVL4 protein were classified as potential orthologues or homologues. Presumed ELOVL4 orthologues and the best-matched sequence homologue from each of a variety of species were selected and aligned using the Clustal W (version 1.8) program (Figure 6.3). A wide range of species including monkey, mouse, rat, insects, fish, fungus, plant, nematode, and yeast varieties were found to contain ELOVL4 homologous protein sequences. True orthologues were identified in monkey and mouse, each with greater than 90% amino acid identity with the human protein. Ninety-one amino acids are strongly evolutionarily conserved, present among eight or more of the homologous sequences examined, and representing almost one-third of the entire protein. These residues tend to lie in or adjacent to putatively functionally significant regions of the protein such as the diiron-oxy-binding HXXHH redox centre motif of the elongase enzyme and the transmembrane domains (Figure 6.3). A total of 72 distinct sequences identified to date in 18 different organisms share at least 23% identity with the human ELOVL4 protein. This significant level of protein sequence similarity between such diverse species is suggestive of a critical cellular role that has been maintained throughout evolution. A phylogenetic tree depicting the evolutionary relationship between these sequences was generated using the Clustal W (version 1.82) and PHYLIP (version 3.5) package of programs (Figure 6.4). Belonging to distinct clades, each of the human proteins identified appears to be more related to gene products from other species than to other members of the human ELO protein family, while many of the putative non-human elongases exhibit greater intraspecies similarity and interspecies divergence. In most cases, the invertebrate ELO proteins group according to species while the vertebrate ELOs span a number of distinct clusters (Figure 6.4).

Despite sharing certain functional motifs including histidine clusters and ER membrane-spanning domains,[32,33] fatty acid desaturases, another class of enzymes involved in the process of fatty acid elongation,[34] were not found to be represented among the list of putative ELO proteins. Furthermore, no significant overall sequence similarity was detected between the human ELOVL4 protein and any of the known human desaturase gene products identified to date. Therefore fatty acid elongases and desaturases represent two independent enzymatic protein families.

6.3.3 Immunodetection of ELOVL4 Protein in Mammalian Retinas

To confirm the existence of ELOVL4 protein homologues in multiple species, polyclonal antibodies raised against a peptide derived from the human sequence were used to probe immunoblots of total protein extracts prepared from the retinas of various animals. A single intense band of approximately 37 kDa was detected in mouse, cow, cat, and rat retinal extracts (Figure 6.5). A similarly prominent band was detected in protein extracts prepared from human retinal tissue. Competition assays performed using the ELOVL4 antibodies in the presence and absence of the peptide antigen verified the specificity of the antibodies for this 37-kDa band, which is consistent with the predicted molecular weight of the human ELOVL4 protein (36.8 kDa). Thus the ELOVL4 protein is present in the retinas of different mammalian species, and the amino acid sequence corresponding to the peptide antigen used for polyclonal antibody production is sufficiently conserved between humans, mice, cows, cats, and rats, in which ELOVL4 protein species of the identical size are detected.

6.3.4 Distribution of ELOVL4 Protein in Mammalian Retinas

The localization of ELOVL4 proteins in specific cell layers within the retinas of various mammalian species was examined by immunohistochemistry. In human retina tissue sections, strong ELOVL4 immunostaining was observed in the photoreceptor layer, particularly within the inner segments of the photoreceptor cells (Figure 6.6A). Uniform staining was detected across the length of this retinal layer in the sections examined, and was found in both rod and cone cells. Slight antibody reactivity was also noted in cell processes extending from the outer plexiform layer. Cell bodies were not observed to contain ELOVL4 protein, as evidenced by the absence of immunostaining in this subcellular region. In mouse, cow, cat, and rat retinal sections, a similar protein localization pattern was observed, with staining above background levels predominantly confined to the photoreceptor inner segment layer (Figures 6.6B through 6.6E). The intensely pigmented RPE cell layer observed in the mouse, cow, and cat sections make it difficult to ascertain if staining due to the ELOVL4 antibodies is present in this region. However, slight RPE staining was detected in sections derived from melanin-deficient albino rats, indicating localization of ELOVL4 proteins to RPE cells (Figure 6.6E). Human retina sections also exhibited increased RPE staining in the presence of the ELOVL4 antiserum (Figure 6.6A), which is in accordance with the detection of ELOVL4 RNA transcripts in human RPE tissue (Figure 6.2). No significant immunostaining was observed in control sections incubated with pre-immune serum instead of the ELOVL4 polyclonal antibodies (Figures 6.6F through 6.6J). The sub-retinal localization of the ELOVL4 protein thus appears to be consistent for a variety of mammalian species.

6.4 DISCUSSION

The studies presented clearly demonstrate that the *ELOVL4* gene and its expression are highly conserved across a diverse array of species. At the genomic DNA, mRNA, and protein levels, and through both *in silico* and classical experimental techniques, similarities in sequence and gene expression profiles among retinal tissues from many different types of animals have consistently been observed. Expression of the *ELOVL4* gene has previously been assessed in human retina and other cell types at the level of mRNA production, however no analysis of ELOVL4 protein has been performed until now. These results confirm those of previous studies that have examined *ELOVL4* gene expression[19,23] and extend their findings.

Consistent with previous reports,[19,23] both the presumed full-length 2.9-kb mRNA transcript and a shorter 2.4-kb transcript were detected in human retinal tissue. However the existence of additional transcripts that bind specifically to *ELOVL4* coding sequences was also repeatedly suggested. Nucleic acid sequence database searches using the hybridization probe sequence as bait do not reveal any other known genes that are predicted to produce the transcript sizes detected. It is possible that these novel RNA species represent the products of genes that have yet to be defined as *ELOVL4* homologues. Multiple paralogues of other genes involved in fatty acid elongation such as desaturases have been identified in various species, giving rise to physically and catalytically distinct enzymes.[35-37] Furthermore, some members of the human ELO gene family share a greater degree of sequence similarity between them than others, suggesting the existence of protein "subfamilies" whose members may exhibit more similar functional characteristics as well. This observation, along with the ongoing discovery of novel human elongase genes (most recently the LCE gene[38,39]), may reveal the existence of multiple *ELOVL4*-like sequences arising from distinct genetic loci. These sequences may give rise to the approximately 4.5-kb and 0.25-kb transcripts appearing on the northern blots. The 2.4-kb mRNA species is thought to arise by alternative polyadenylation of the nascent *ELOVL4* RNA transcript.[23] Alternative promoter usage is another means by which different transcripts can be produced from a single genetic locus. Thus both pre-transcriptional and post-transcriptional processing of as yet unidentified *ELOVL4*-like genes and gene products may account for the results of northern analysis observed. As these novel transcripts are detected in the retina of different species, the mechanism by which they arise appears also to be conserved. All four of the transcripts are seen in tissue derived from both the central and peripheral retina, however there is more limited gene expression noted in other human ocular tissues. This finding resembles the expression profile of other retinal genes,[40,41] and indicates that retina-specific gene function may be imparted through differential rather than retina-specific gene expression.

Bioinformatic analysis reveals that ELOVL4 protein orthologues have been identified thus far in monkey[42] and in mouse,[19] exhibiting 97% and 93% amino acid identity, respectively, with the human protein. In addition, a wide variety of evolutionarily divergent species also harbour ELOVL4-homologous sequences, suggesting the importance of this gene or similar entities to cellular function. As expected, regions of the protein with the highest degree of conservation coincide with predicted functional motifs. Highly conserved amino acid residues also exist outside of these domains and indicate the significance of other protein regions that have yet to be defined within ELOVL4, such as potential substrate-binding, protein-protein interaction, and regulatory domains.

The ELOVL4 protein is a member of a large family of known and putative fatty acid elongases, sharing from 23% to 97% amino acid sequence identity with 72 different proteins identified in this family to date. Results of phylogenetic analysis indicate that, in general, the intraspecies divergence of vertebrate ELO proteins is greater than that of invertebrate ELOs. Proteins from moss, fungus, plant, nematode, and single-celled organisms cluster according to species, while mammalian ELO proteins tend to cluster according to protein family membership instead of species designations. This is evident for all of the human ELO protein sequences examined, including ELOVL4, and suggests that the evolutionary divergence of ELOs into 'subfamilies' occurred before the divergence of individual vertebrate species. This in turn supports the hypothesis that the genes encoding these proteins have evolved from common ancestral genes that play distinct roles in fatty acid biosynthesis and are part of a highly conserved metabolic mechanism.

The exclusion of fatty acid desaturases from the extensive list of ELOVL4 protein homologues identified indicates that physically distinct enzymes are responsible for catalyzing different reactions in the fatty acid elongation process in many species. Clearly the HXXHH motifs, transmembrane domains, and putative ER-targeting signals are not sufficient to impart significant overall amino acid sequence homology between members of the elongase and desaturase protein families, as might be expected for enzymes that are involved in distinct catalytic activities and that have requirements for recognizing chemically different substrates. This contrasts with other lipid biosynthetic processes such as *de novo* fatty acid synthesis, in which the fatty acid synthase enzyme consists of multifunctional polypeptides in animal cells but distinct proteins in bacteria.[43] However it is believed that the reactions constituting the fatty acid desaturation and chain elongation pathway are catalyzed by a closely coordinated or concerted enzyme system.[44]

Towards the aim of further studying the ELOVL4 protein, this report describes for the first time the generation and use of polyclonal antibodies that specifically recognize the expected 37-kDa ELOVL4 protein in retinal extracts from different species. Immunohistochemical studies performed with these antibodies reveals conserved sub-retinal localization of the protein in photoreceptor cells, and more specifically within the inner segments of both rods and cones. This finding is consistent with the results of *in situ* hybridization experiments previously reported.[19] The cellular biosynthetic machinery is located in the inner segments of photoreceptor cells.[45] Accordingly, enzymes involved in various biosynthetic processes in the retina have been localized to this subretinal layer.[46-50] Thus it is expected that ELOVL4, an enzyme putatively involved in lipid synthesis, is also found in this subcellular location. Furthermore, the process of fatty acid elongation in which the ELOVL4 protein is thought to function primarily occurs at the endoplasmic reticulum membrane.[34,51] This coupled with the putative ER-localization motif at the carboxyl terminus of the protein sequence predicts that ELOVL4 would be found in the ER, which in turn is located in photoreceptor inner segments.[52] The disease mutations thus far identified in the *ELOVL4* gene in STGD3/adMD-affected pedigrees are believed to lead to the production of a prematurely truncated protein that lacks the ER-localization signal, rendering the enzyme dysfunctional as a result of aberrant protein targeting. Photoreceptor inner segment labeling by ELOVL4 antibodies described here is the first demonstration of subcellular localization of this protein and is consistent with its putative location in the endoplasmic reticulum within cells.

Several of the disease genes that underlie atrophic macular degeneration phenotypes are expressed in both rod and cone photoreceptors,[53-55] although the cone-rich macular region is the primary target in these conditions.[8,56,57] It is interesting to note that the *ELOVL4* gene is also expressed in both types of photoreceptor cells. The *ABCA4*, *RDS*, and *RPGR* genes, whose protein products have been detected in both rods and cones, are also mutated in disorders affecting the peripheral retina.[58-60] However, no disease-associated sequence alterations in *ELOVL4* have as yet been reported in patients with peripheral retina disorders despite mutational analysis of the gene within such individuals.[61,62] Elucidation of the pathological mechanisms by which mutation of each of these genes leads to the development of diseases affecting different regions of the retina awaits further study.

It has been hypothesized that ELOVL4 may be an enzyme responsible for catalyzing an elongation step in the synthesis of docosahexaenoic acid (DHA),[19] the most abundant polyunsaturated fatty acid species in vertebrate retinal lipids.[62] It is also possible that this protein is involved in the synthesis of other essential fatty acids whose deficiency may lead to impaired retinal function. Alteration of the lipid environment of retinal cells has been shown to influence retinal protein function, photoreceptor cell development and cell survival.[63-68] Thus dysfunction of a gene that results in aberrant lipid synthesis in the retina is predicted to have deleterious effects and may account for the pathogenesis observed in cases of STGD3/adMD and other retinal dystrophies in which *ELOVL4* mutations are identified. Continued characterization of the physical and functional properties of ELOVL4 will provide valuable insight towards unraveling its precise biological role in the retina, which in turn may lead to the development of potential therapeutic strategies for individuals suffering from the ocular disorders arising from mutations in the *ELOVL4* gene.

6.5 ACKNOWLEDGEMENTS

The authors thank Janet Wagner for providing archival human tissue samples, Mitchell S. Gillett for preparation of animal tissue sections, Austra Liepa for animal care, Miguel A. Cabrita for assistance with the phylogenetic analysis, and Dushanka Deretic for technical advice.

6.6 REFERENCES

1. Hadden OB, Gass JD. (1976) Fundus flavimaculatus and Stargardt's disease. *Am. J. Ophthalmol.* 82:527-539.

2. Noble KG, Carr RE. (1979) Stargardt's disease and fundus flavimaculatus. *Arch. Ophthalmol.* 97:1281-1285.

3. Eagle RC Jr, Lucier AC, Bernardino VB Jr, Yanoff M. (1980) Retinal pigment epithelial abnormalities in fundus flavimaculatus: a light and electron microscopic study. *Ophthalmology* 87:1189-1200.

4. Lopez PF, Maumenee IH, de la Cruz Z, Green WR. (1990) Autosomal-dominant fundus flavimaculatus. Clinicopathologic correlation. *Ophthalmology* 97:798-809.

5. Lois N, Holder GE, Bunce C, Fitzke FW, Bird AC. (2001) Phenotypic subtypes of Stargardt macular dystrophy-fundus flavimaculatus. *Arch. Ophthalmol.* 119:359-369.

6. Donoso LA, Edwards AO, Frost A, Vrabec T, Stone EM, Hageman GS, Perski T. (2001) Autosomal dominant Stargardt-like macular dystrophy. *Surv. Ophthalmol.* 46:149-163.

7. Kaplan J, Bonneau D, Frezal J, Munnich A, Dufier JL. (1990) Clinical and genetic heterogeneity in retinitis pigmentosa. *Hum. Genet.* 85:635-642.

8. Allikmets R, Singh N, Sun H, Shroyer NF, Hutchinson A, Chidambaram A, Gerrard B, Baird L, Stauffer D, Peiffer A, Rattner A, Smallwood P, Li Y, Anderson KL, Lewis RA, Nathans J, Leppert M, Dean M, Lupski JR. (1997) A photoreceptor cell-specific ATP-binding transporter gene (*ABCR*) is mutated in recessive Stargardt macular dystrophy. *Nat .Genet.* 15:236-246.

9. Daiger SP. RetNet, the Retinal Information Network http://www.sph.uth.tmc.edu/Retnet/ .

10. Sun H, Molday RS, Nathans J. (1999) Retinal stimulates ATP hydrolysis by purified and reconstituted ABCR, the photoreceptor-specific ATP-binding cassette transporter responsible for Stargardt disease. *J. Biol. Chem.* 274:8269-8281.

11. Weng J, Mata NL, Azarian SM, Tzekov RT, Birch DG, Travis GH. (1999) Insights into the function of Rim protein in photoreceptors and etiology of Stargardt's disease from the phenotype in *abcr* knockout mice. *Cell* 98:13-23.

12. van Driel MA, Maugeri A, Klevering BJ, Hoyng CB, Cremers FP. (1998) *ABCR* unites what ophthalmologists divide(s). *Ophthalmic Genet.* 19:117-122.

13. Shroyer NF, Lewis RA, Allikmets R, Singh N, Dean M, Leppert M, Lupski JR. (1999) The rod photoreceptor ATP-binding cassette transporter gene, *ABCR*, and retinal disease: from monogenic to multifactorial. *Vision Res.* 39:2537-2544.

14. Stone EM, Nichols BE, Kimura AE, Weingeist TA, Drack A, Sheffield VC. (1994) Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. *Arch. Ophthalmol.* 112:765-772.

15. Griesinger IB, Sieving PA, Ayyagari R. (2000) Autosomal dominant macular atrophy at 6q14 excludes CORD7 and MCDR1/PBCRA loci. *Invest. Ophthalmol. Vis. Sci.* 41:248-255.

16. Kniazeva M, Chiang MF, Morgan B, Anduze AL, Zack DJ, Han M, Zhang K. (1999) A new locus for autosomal dominant stargardt-like disease maps to chromosome 4. *Am. J. Hum. Genet.* 64:1394-1399.

17. Lagali PS, MacDonald IM, Griesinger IB, Chambers ML, Ayyagari R, Wong PW. (2000) Autosomal dominant Stargardt-like macular dystrophy segregating in a large Canadian family. *Can. J. Ophthalmol.* 35:315-324.

18. Gass JDM. (1987) Stargardt's disease (fundus flavimaculatus). In: Gass JDM, ed. *Stereoscopic Atlas of Macular Diseases: Diagnosis and Treatment*. St. Louis: Mosby; pp.256-261.

19. Zhang K, Kniazeva M, Han M, Li W, Yu Z, Yang Z, Li Y, Metzker ML, Allikmets R, Zack DJ, Kakuk LE, Lagali PS, Wong PW, MacDonald IM, Sieving PA, Figueroa DJ, Austin CP, Gould RJ, Ayyagari R, Petrukhin K. (2001) A 5-bp deletion in *ELOVL4* is associated with two related forms of autosomal dominant macular dystrophy. *Nat. Genet.* 27:89-93.

20. Toke DA, Martin CE. (1996) Isolation and characterization of a gene affecting fatty acid elongation in Saccharomyces cerevisiae. *J. Biol. Chem.* 271:18413-18422.

21. Oh CS, Toke DA, Mandala S, Martin CE. (1997) *ELO2* and *ELO3*, homologues of the *Saccharomyces cerevisiae ELO1* gene, function in fatty acid elongation and are required for sphingolipid formation. *J. Biol. Chem.* 272:17376-17384.

22. Jackson MR, Nilsson T, Peterson PA. (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO J.* 9:3153-3162.

23. Edwards AO, Donoso LA, Ritter R 3rd. (2001) A novel gene for autosomal dominant Stargardt-like macular dystrophy with homology to the SUR4 protein family. *Invest. Ophthalmol. Vis. Sci.* 42:2652-2663.

24. Vrabec TR, Tantri A, Edwards A, Frost A, Donoso LA. (2003) Autosomal dominant Stargardt-like macular dystrophy: identification of a new family with a mutation in the *ELOVL4* gene. *Am. J. Ophthalmol.* 136:542-545.

25. Bernstein PS, Tammur J, Singh N, Hutchinson A, Dixon M, Pappas CM, Zabriskie NA, Zhang K, Petrukhin K, Leppert M, Allikmets R. (2001) Diverse macular dystrophy phenotype caused by a novel complex mutation in the *ELOVL4* gene. *Invest. Ophthalmol. Vis. Sci.* 42:3331-3336.

26. Bernstein SL, Wong P. (1998) Regional expression of disease-related genes in human and monkey retina. *Mol. Vis.* 4:24.

27. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.

28. Thompson JD, Higgins DG, Gibson TJ. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.

29. Felsenstein J. (1988) Phylogenies from molecular sequences: inference and reliability. *Annu. Rev. Genet.* 22:521-565.

30. Felsenstein J. (1989) PHYLIP -- Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164-166.

31. Finley JC, Grossman GH, Dimeo P, Petrusz P. (1978) Somatostatin-containing neurons in the rat brain: widespread distribution revealed by immunocytochemistry after pretreatment with pronase. *Am. J. Anat.* 153:483-488.

32. Tocher DR, Leaver MJ, Hodgson PA. (1998) Recent advances in the biochemistry and molecular biology of fatty acyl desaturases. *Prog. Lipid Res.* 37:73-117.

33. Shanklin J, Cahoon EB. (1998) Desaturation and related modifications of fatty acids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49:611-641.

34. Sprecher H. (1999) An update on the pathways of polyunsaturated fatty acid metabolism. *Curr. Opin. Clin. Nutr. Metab. Care* 2:135-138.

35. MacKenzie DA, Carter AT, Wongwathanarat P, Eagles J, Salt J, Archer DB. (2002) A third fatty acid Δ^9-desaturase from *Mortierella alpina* with a different substrate specificity to ole1p and ole2p. *Microbiology* 148:1725-1735.

36. Tabor DE, Xia YR, Mehrabian M, Edwards PA, Lusis AJ. (1998) A cluster of stearoyl CoA desaturase genes, *Scd1* and *Scd2*, on mouse chromosome 19. *Mamm. Genome* 9:341-342.

37. Zheng Y, Prouty SM, Harmon A, Sundberg JP, Stenn KS, Parimoo S. (2001) *Scd3*–a novel gene of the stearoyl-CoA desaturase family with restricted expression in skin. *Genomics* 71:182-191.

38. Moon Y-A, Shah NA, Mohapatra S, Warrington JA, Horton JD. (2001) Identification of a mammalian long chain fatty acyl elongase regulated by sterol regulatory element-binding proteins. *J. Biol. Chem.* 276:45358-45366.

39. Matsuzaka T, Shimano H, Yahagi N, Yoshikawa T, Amemiya-Kudo M, Hasty AH, Okazaki H, Tamura Y, Iizuka Y, Ohashi K, Osuga J, Takahashi A, Yato S, Sone H, Ishibashi S, Yamada N. (2002) Cloning and characterization of a mammalian fatty

acyl-CoA elongase as a lipogenic enzyme regulated by SREBPs. *J. Lipid Res.* 43:911-920.

40. Lagali PS, Ayyagari R, Kakuk LE, MacDonald IM, Wong PW. (2000) Retinal pigment epithelial-expressed genes as putative candidates for inherited retinal dystrophies mapping to chromosome 6q. *Am. J. Hum. Genet.* 67 Suppl.2:391.

41. Lagali PS, Kakuk LE, Griesinger IB, Wong PW, Ayyagari R. (2002) Identification and characterization of *C6orf37*, a novel candidate human retinal disease gene on chromosome 6q14. *Biochem. Biophys. Res. Commun.* 293:356-365.

42. Umeda S, Ayyagari R, Suzuki MT, Ono F, Iwata F, Fujiki K, Kanai A, Takada Y, Yoshikawa Y, Tanaka Y, Iwata T. (2003) Molecular cloning of *ELOVL4* gene from cynomolgus monkey (*Macaca fascicularis*). *Exp. Anim.* 52:129-135.

43. Wakil SJ, Stoops JK, Joshi VC. (1983) Fatty acid synthesis and its regulation. *Annu. Rev. Biochem.* 52:537-579.

44. Cook HW, Clarke JTR, Spence MW. (1983) Concerted stimulation and inhibition of desaturation, chain elongation, and esterification of essential fatty acids by cultured neuroblastoma cells. *J. Biol. Chem.* 258:7586-7591.

45. Oyster CW. (1999) *The Human Eye: Structure and Function*. Sunderland, MA: Sinauer Associates; pp.585-586.

46. Wiechmann AF, Bok D, Horwitz J. (1985) Localization of hydroxyindole-O-methyltransferase in the mammalian pineal gland and retina. *Invest. Ophthalmol. Vis. Sci.* 26:253-265.

47. Nguyen-Legros J, Krieger M, Simon A. (1994) Immunohistochemical localization of L-dopa and aromatic L-amino acid-decarboxylase in the rat retina. *Invest. Ophthalmol. Vis. Sci.* 35:2906-2915.

48. Fischer AJ, Wallman J, Mertz JR, Stell WK. (1999) Localization of retinoid binding proteins, retinoid receptors, and retinaldehyde dehydrogenase in the chick eye. *J. Neurocytol.* 28:597-609.

49. López-Costa JJ, Goldstein J, Saavedra, JP. (1997) Neuronal and macrophagic nitric oxide synthase isoforms distribution in normal rat retina. *Neurosci. Lett.* 232:155-158.

50. Shin DH, Lim HS, Cho SK, Lee HY, Lee HW, Lee KH, Chung YH, Cho SS, Ik Cha C, Hwang DH. (2000) Immunocytochemical localization of neuronal and inducible nitric oxide synthase in the retina of zebrafish, *Brachydanio rerio*. *Neurosci. Lett.* 292:220-222.

51. Cinti DL, Cook L, Nagi MN, Suneja SK. (1992) The fatty acid chain elongation system of mammalian endoplasmic reticulum. *Prog. Lipid Res.* 31:1-51.

52. Young RW. (1967) The renewal of photoreceptor cell outer segments. *J. Cell. Biol.* 33:61-72.

53. Molday LL, Rabin AR, Molday RS. (2000) ABCR expression in foveal cone photoreceptors and its role in Stargardt macular dystrophy. *Nat. Genet.* 25:257-258.

54. Arikawa K, Molday LL, Molday RS, Williams DS. (1992) Localization of peripherin/rds in the disk membranes of cone and rod photoreceptors: relationship to disk membrane morphogenesis and retinal degeneration. *J. Cell. Biol.* 116:659-667.

55. Hong DH, Pawlyk BS, Shang J, Sandberg MA, Berson EL, Li T. (2000) A retinitis pigmentosa GTPase regulator (RPGR)-deficient mouse model for X-linked retinitis pigmentosa (RP3). *Proc. Natl. Acad. Sci. U. S. A.* 97:3649-3654.

56. Wells J, Wroblewski J, Keen J, Inglehearn C, Jubb C, Eckstein A, Jay M, Arden G, Bhattacharya S, Fitzke F, Bird A. (1993) Mutations in the human retinal degeneration slow (*RDS*) gene can cause either retinitis pigmentosa or macular dystrophy. *Nat. Genet.* 3:213-218.

57. Ayyagari R, Demirci F, Liu J, Bingham EL, Stringham H, Kakuk LE, Boehnke M, Gorin MB, Richards JE, Sieving PA. (2002) X-Linked Recessive Atrophic Macular Degeneration from *RPGR* Mutation. *Genomics* 80:166-171.

58. Martinez-Mir A, Paloma E, Allikmets R, Ayuso C, del Rio T, Dean M, Vilageliu L, Gonzalez-Duarte R, Balcells S. (1998) Retinitis pigmentosa caused by a homozygous mutation in the Stargardt disease gene *ABCR*. *Nat. Genet.* 18:11-12.

59. Kajiwara K, Hahn LB, Mukai S, Travis GH, Berson EL, Dryja TP. (1991) Mutations in the human retinal degeneration slow gene in autosomal dominant retinitis pigmentosa. *Nature* 354:480-483.

60. Meindl A, Dry K, Herrmann K, Manson F, Ciccodicola A, Edgar A, Carvalho MR, Achatz H, Hellebrand H, Lennon A, Migliaccio C, Porter K, Zrenner E, Bird A, Jay M, Lorenz B, Wittwer B, D'Urso M, Meitinger T, Wright A. (1996) A gene (*RPGR*) with homology to the RCC1 guanine nucleotide exchange factor is mutated in X-linked retinitis pigmentosa (RP3). *Nat. Genet.* 13:35-42.

61. Li Y, Marcos I, Borrego S, Yu Z, Zhang K, Antiñolo G. (2001) Evaluation of the *ELOVL4* gene in families with retinitis pigmentosa linked to the *RP25* locus. *J. Med. Genet.* 38:478-480.

62. Rivolta C, Ayyagari R, Sieving PA, Berson EL, Dryja TP. (2003) Evaluation of the *ELOVL4* gene in patients with autosomal recessive retinitis pigmentosa and Leber congenital amaurosis. *Mol. Vis.* 9:49-51.

63. Futterman S, Andrews JS. (1964) The fatty acid composition of human retinal vitamin A ester and the lipids of human retinal tissue . *Invest. Ophthalmol.* 3:441-444.

64. Brown MF. (1994) Modulation of rhodopsin function by properties of the membrane bilayer. *Chem. Phys. Lipids* 73:159-180.

65. Ahn J, Wong JT, Molday RS. (2000) The effect of lipid environment and retinoids on the ATPase activity of ABCR, the photoreceptor ABC transporter responsible for Stargardt macular dystrophy. *J. Biol. Chem.* 275:20399-20405.

66. Suh M, Wierzbicki AA, Lien EL, Clandinin MT. (2000) Dietary 20:4n-6 and 22:6n-3 modulates the profile of long- and very-long-chain fatty acids, rhodopsin content, and kinetics in developing photoreceptor cells. *Pediatr. Res.* 48:524-530.

67. Rotstein NP, Aveldano MI, Barrantes FJ, Roccamo AM, Politi LE. (1997) Apoptosis of retinal photoreceptors during development in vitro: protective effect of docosahexaenoic acid. *J. Neurochem.* 69:504-513.

68. Rotstein NP, Aveldano MI, Barrantes FJ, Politi LE. (1996) Docosahexaenoic acid is required for the survival of rat retinal photoreceptors *in vitro*. *J. Neurochem.* 66:1851-1959.

69. Mizota A, Sato E, Taniai M, Adachi-Usami E, Nishikawa M. (2001) Protective effects of dietary docosahexaenoic acid against kainate-induced retinal degeneration in rats. *Invest. Ophthalmol. Vis. Sci.* 42:216-221.

Figure 6.1 Genomic sequence conservation of the *ELOVL4* gene in diverse species. Southern zooblot analysis of genomic DNA derived from various species demonstrates evolutionary conservation of the *ELOVL4* gene. Human, monkey, rat, and mouse DNA samples were analyzed with a radiolabeled probe corresponding to the human *ELOVL4* gene exon 6 sequence, while an *ELOVL4* cDNA fragment was used to probe the dog, cow, rabbit, chicken, and yeast DNA samples.

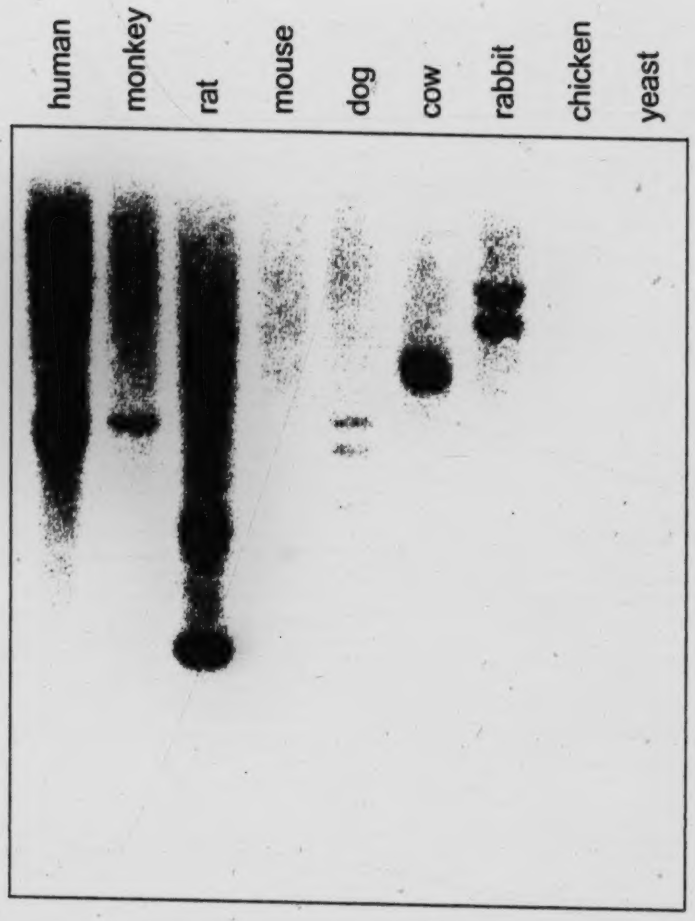

human
monkey
rat
mouse
dog
cow
rabbit
chicken
yeast

Figure 6.2 Ocular *ELOVL4* gene expression in mammalian species. Northern blot analysis of human, monkey, mouse, cow, and pig ocular tissues demonstrates differential expression of ELOVL4 RNA transcripts between different regions of the eye and between primate and mammalian species. Each lane contains 5 μg of total RNA extracted from the indicated tissues. Four distinct transcripts are detected upon probing with the ELOVL4 cDNA fragment, indicated by arrowheads to the right of the blots. Results of analysis of the blots with a human β-actin control probe are shown in the bottom panels. Positions of RNA size markers are shown at left (*kb*, kilobases).

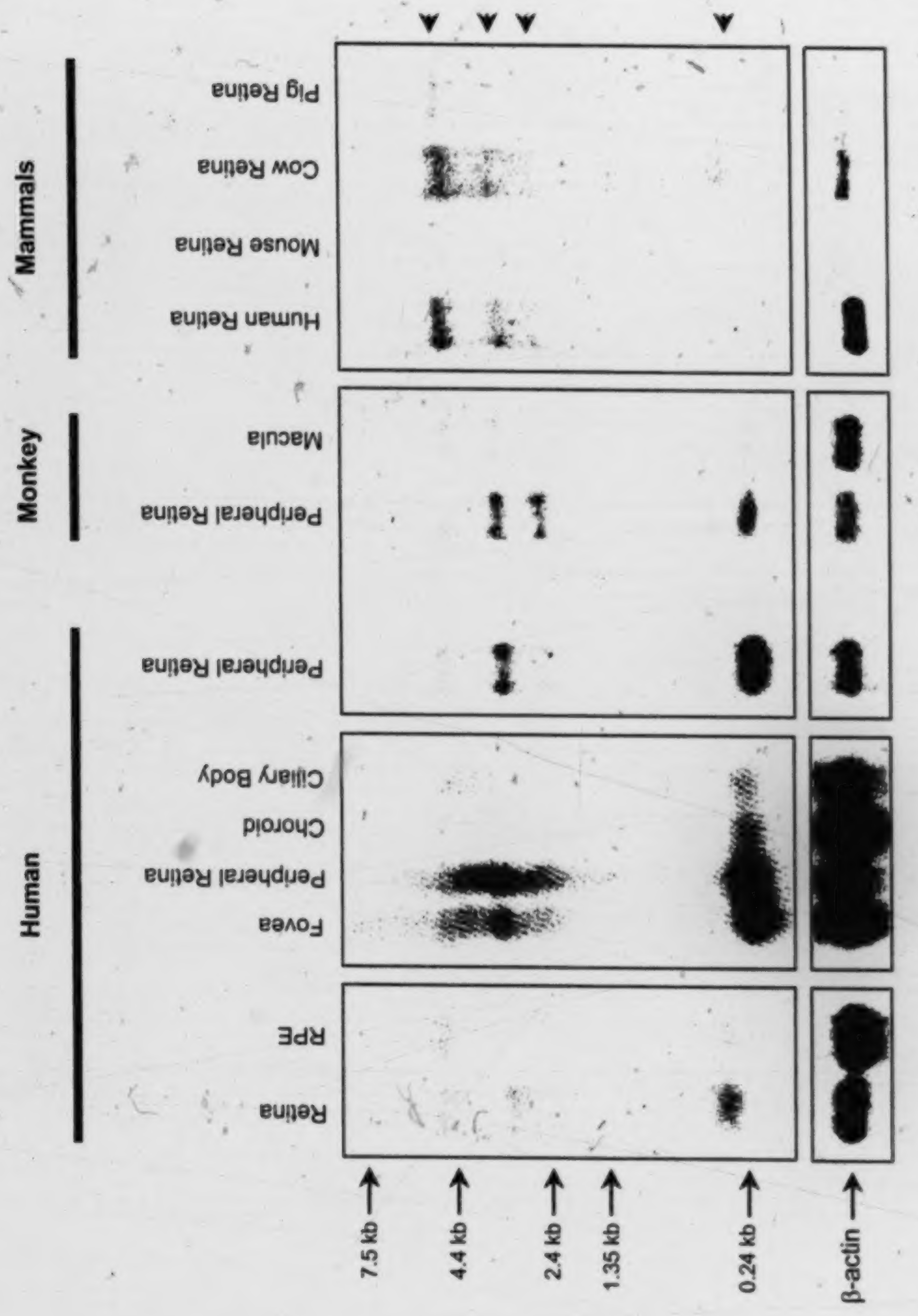
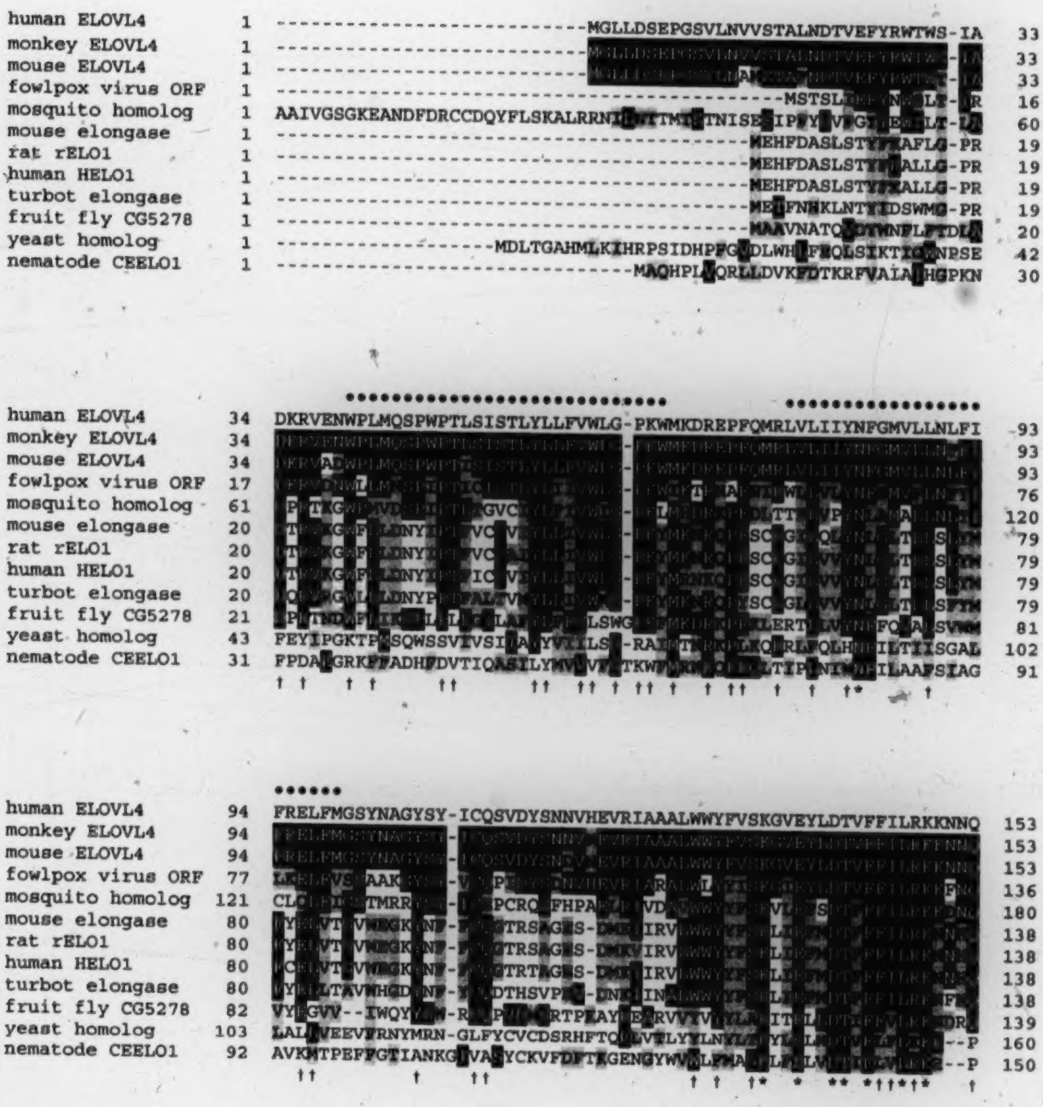

Human
Monkey
Mammals
Retina
RPE
Fovea
Peripheral Retina
Choroid
Ciliary Body
Peripheral Retina
Peripheral Retina
Macula
Human Retina
Mouse Retina
Cow Retina
Pig Retina
7.5 kb
4.4 kb
2.4 kb
1.35 kb
0.24 kb
β-actin

Figure 6.3 Sequence alignment of the human ELOVL4 protein with known orthologues and most similar homologues identified in other species. Monkey and mouse ELOVL4 orthologues are labeled, while the remaining sequences represent proteins identified to date that exhibit the greatest degree of homology with human ELOVL4 in each of the organisms listed. Amino acid sequences were aligned using the Clustal W program (version 1.8). Hyphens represent gaps introduced to maintain the alignment. The amino acid residues are numbered to the left and right of the sequences. Identical residues among the sequences in the alignment are shown on a black background, and conservative amino acid substitutions are shaded in grey. Residues identical among all twelve sequences are marked with asterisks, while amino acids common to eight or more sequences are marked with a cross. Predicted transmembrane domains in human ELOVL4 are indicated with a series of dots above each of the corresponding hydrophobic stretches of amino acids. The diiron-oxo-binding 'histidine box' motif shared by fatty acid elongase enzymes is shown in boldface. GenBank accession numbers for each protein sequence listed are as follows: human (*H. sapiens*) ELOVL4, AAG47668; monkey (*M. fascicularis*) ELOVL4, Q95K73; mouse (*M. musculus*) ELOVL4, AAG47667; fowlpox virus ORF FPV048, AAF44392; mosquito (*A. gambiae*) homologue, EAA07022; mouse (*M. musculus*) elongase, NP_599016; rat (*R. norvegicus*) rELO1, BAB69887; human (*H. sapiens*) HELO1, AAF70631; turbot (*S. maximus*) elongase, AAL69984; fruit fly (*D. melanogaster*) CG5278 gene product, AAF56019; yeast (*S. pombe*) homologue, T50139; nematode (*C. elegans*) CEELO1, AAF70462.


human ELOVL4 1 MGLLDSEPGSVLNVVSTALNDTVEFYRWTWS-IA 33
monkey ELOVL4 1 33
mouse ELOVL4 1 33
fowlpox virus ORF 1 -MSTSL 16
mosquito homolog 1 AAIVGSGKEANDFDRCCDQYFLSKALRRN 60
mouse elongase 1 MEHFDASLSTYFKAFLG-PR 19
rat rELO1 1 MEHFDASLSTYF ALLG-PR 19
human HELO1 1 MEHFDASLSTYFKALLG-PR 19
turbot elongase 1 MEEFNHKLNTYIDSWMG-PR 19
fruit fly CG5278 1 MAAVNATQ 20
yeast homolog 1 MDLTGAHMLKIHRPSIDHPFG 42
nematode CEELO1 1 MAQHPL 30
human ELOVL4 34 DKRVENWPLMQSPWPTLSISTLYLLFVWLG-PKWMKDREPFQMRLVLIIYNFGMVLLNLFI 93
monkey ELOVL4 34 93
mouse ELOVL4 34 93
fowlpox virus ORF 17 76
mosquito homolog 61 120
mouse elongase 20 79
rat rELO1 20 79
human HELO1 20 79
turbot elongase 20 79
fruit fly CG5278 21 81
yeast homolog 43 FEYIPGKTP SQWSSVIVSI 102
nematode CEELO1 31 FPDA GRK FADH DVTIQASILYMV 91
human ELOVL4 94 FRELFMGSYNAGYSY-ICQSVDYSNNVHEVRIAAALWWYFVSKGVEYLDTVFFILRKKNNQ 153
monkey ELOVL4 94 153
mouse ELOVL4 94 153
fowlpox virus ORF 77 136
mosquito homolog 121 180
mouse elongase 80 138
rat rELO1 80 138
human HELO1 80 138
turbot elongase 80 138
fruit fly CG5278 82 139
yeast homolog 103 LAL VEEV RNYMRN-GLFYCVCDSRHFTQ 160
nematode CEELO1 92 AVK TPEF PGTIANKG VA YCKVFDFT GENGYWV --P 150

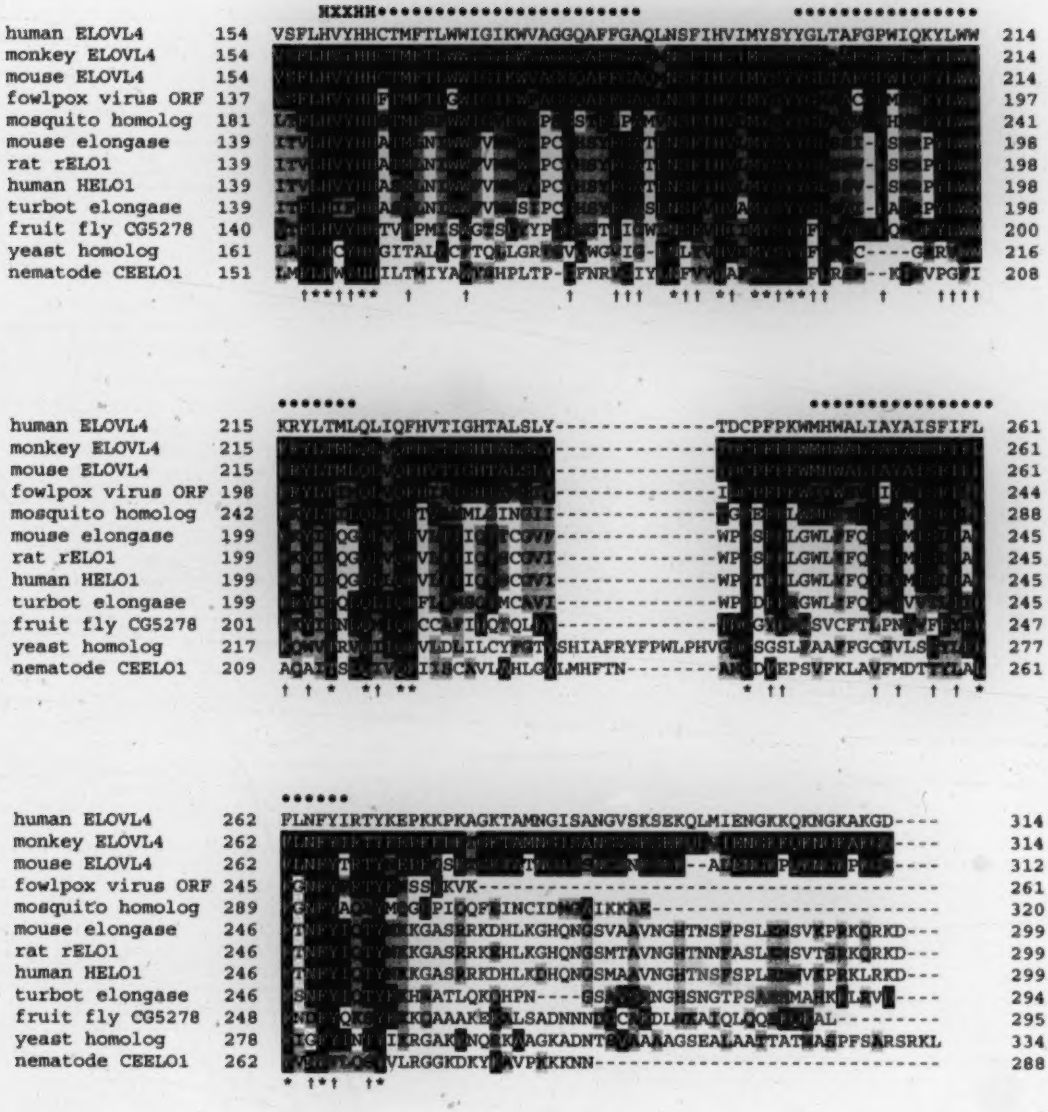

HXXHH
human ELOVL4
monkey ELOVL4
mouse ELOVL4
fowlpox virus ORF
mosquito homolog
mouse elongase
rat rELO1
human HELO1
turbot elongase
fruit fly CG5278
yeast homolog
nematode CEELO1
VSFLHVYHHCTMFTLWWIGIKWVAGGQAFFGAQLNSFIHVIMYSYYGLTAFGPWIQKYLWW
KRYLTMLQLIQFHVTIGHTALSLY-------------TDCPFPKWMHWALIAYAISFIFL
FLNFYIRTYKEPKKPKAGKTAMNGISANGVSKSEKQLMIENGKKQKNGKAKGD----

Figure 6.4 Phylogenetic tree of ELO protein family members. Evolutionary relationships between all known and predicted ELOVL4 protein homologues found in the GenBank protein sequence database to date are depicted. Two- or three-letter species abbreviations precede the protein name or GenBank accession number for each sequence represented. Percentage identity between each sequence and human ELOVL4 (boldface) is indicated in parentheses following each name or accession number. The rooted phenogram shown was generated by aligning amino acid sequences using the Clustal W program (version 1.82) followed by phylogenetic analysis with the PHYLIP software package (version 3.5). Species abbreviations: *Hs*, Homo sapiens (human); *Mf*, Macaca fascicularis (monkey); *Mm*, Mus musculus (mouse); FPV, fowlpox virus; *Ag*, Anopheles gambiae (mosquito); *Rn*, Rattus norvegicus (rat); *Sm*, Scophthalmus maximus (turbot); *Dm*, Drosophila melanogaster (fruit fly); *Pp*, Physcomitrella patens (moss); *Ma*, Mortierella alpina (fungus); *Sp*, Schizosaccharomyces pombe (fission yeast); *Sc*, Saccharomyces cerevisiae (budding yeast); *At*, Arabidopsis thaliana (plant); *Os*, Oryza sativa (rice); *Dd*, Dictyostelium discoideum (amoebae); *Ce*, Caenorhabditis elegans (nematode); *Lm*, Leishmania major (trypanosome); *Ig*, Isochrysis galbana (micro-algae).

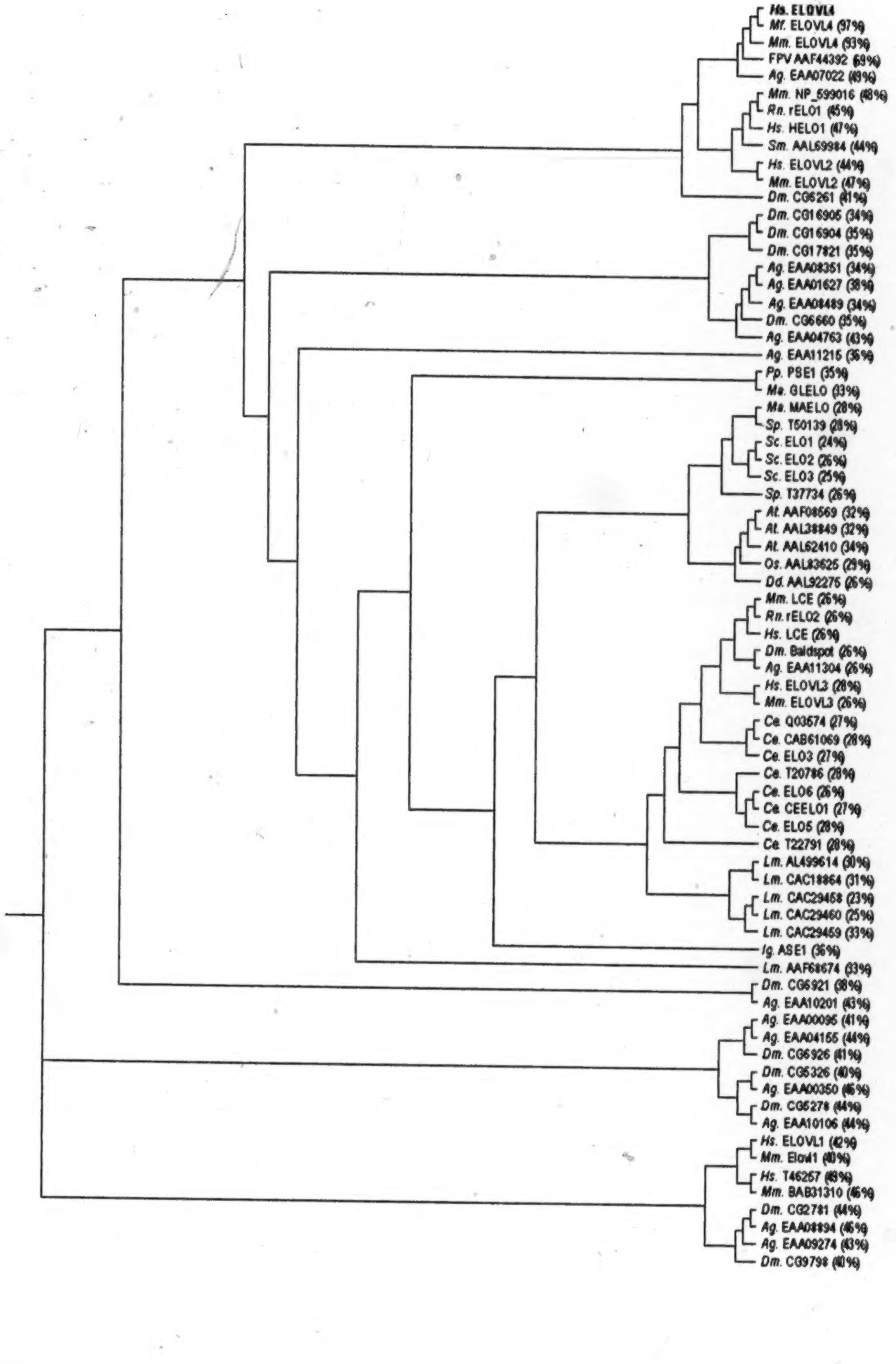

Hs. ELOVL4
Mf. ELOVL4 (97%)
Mm. ELOVL4 (93%)
FPV AAF44392 (69%)
Ag. EAA07022 (48%)
Mm. NP_599016 (48%)
Rn. rELO1 (45%)
Hs. HELO1 (47%)
Sm. AAL69984 (44%)
Hs. ELOVL2 (44%)
Mm. ELOVL2 (47%)
Dm. CG6261 (41%)
Dm. CG16905 (34%)
Dm. CG16904 (35%)
Dm. CG17821 (35%)
Ag. EAA08351 (34%)
Ag. EAA01627 (38%)
Ag. EAA08489 (34%)
Dm. CG6660 (35%)
Ag. EAA04763 (43%)
Ag. EAA11215 (36%)
Pp. PSE1 (35%)
Ma. GLELO (33%)
Ma. MAELO (28%)
Sp. T50139 (28%)
Sc. ELO1 (24%)
Sc. ELO2 (26%)
Sc. ELO3 (25%)
Sp. T37734 (26%)
At. AAF08569 (32%)
At. AAL38849 (32%)
At. AAL62410 (34%)
Os. AAL83625 (29%)
Dd. AAL92275 (26%)
Mm. LCE (26%)
Rn. rELO2 (26%)
Hs. LCE (26%)
Dm. Baldspot (26%)
Ag. EAA11304 (26%)
Hs. ELOVL3 (28%)
Mm. ELOVL3 (26%)
Ce. Q03574 (27%)
Ce. CAB61069 (28%)
Ce. ELO3 (27%)
Ce. T20786 (28%)
Ce. ELO6 (26%)
Ce. CEELO1 (27%)
Ce. ELO5 (28%)
Ce. T22791 (28%)
Lm. AL499614 (30%)
Lm. CAC18864 (31%)
Lm. CAC29458 (23%)
Lm. CAC29460 (25%)
Lm. CAC29459 (33%)
Ig. ASE1 (36%)
Lm. AAF68674 (33%)
Dm. CG6921 (38%)
Ag. EAA10201 (43%)
Ag. EAA00095 (41%)
Ag. EAA04165 (44%)
Dm. CG6926 (41%)
Dm. CG5326 (40%)
Ag. EAA00350 (45%)
Dm. CG5278 (44%)
Ag. EAA10106 (44%)
Hs. ELOVL1 (42%)
Mm. Elovl1 (40%)
Hs. T46257 (43%)
Mm. BAB31310 (45%)
Dm. CG2781 (44%)
Ag. EAA08894 (45%)
Ag. EAA09274 (43%)
Dm. CG9798 (40%)

Figure 6.5 Conservation of ELOVL4 protein synthesis in mammalian retinas. Immunoblots of retinal protein extracts derived from various mammalian species were probed with polyclonal anti-ELOVL4 antiserum, demonstrating conserved ELOVL4 protein production. The 37-kilodalton (kDa) protein corresponding to the predicted size of ELOVL4 is detected in mouse, cow, cat, rat, and human retina tissues. Significant reduction in the intensity of the 37-kDa band is observed in human extracts when blots are incubated with primary antibodies in the presence of excess oompeting peptide antigen ('+ peptide' lane), demonstrating specificity of the ELOVL4 antibodies. Autoradiographic images provided by Ms. Laura E. Kakuk (University of Michigan).

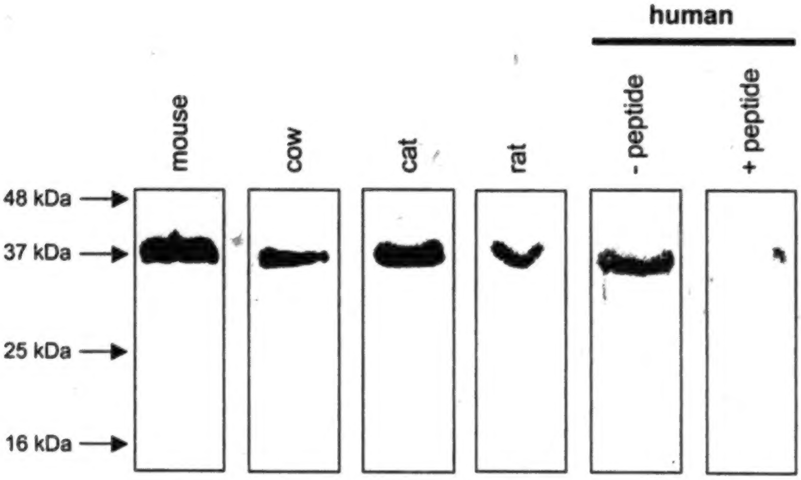
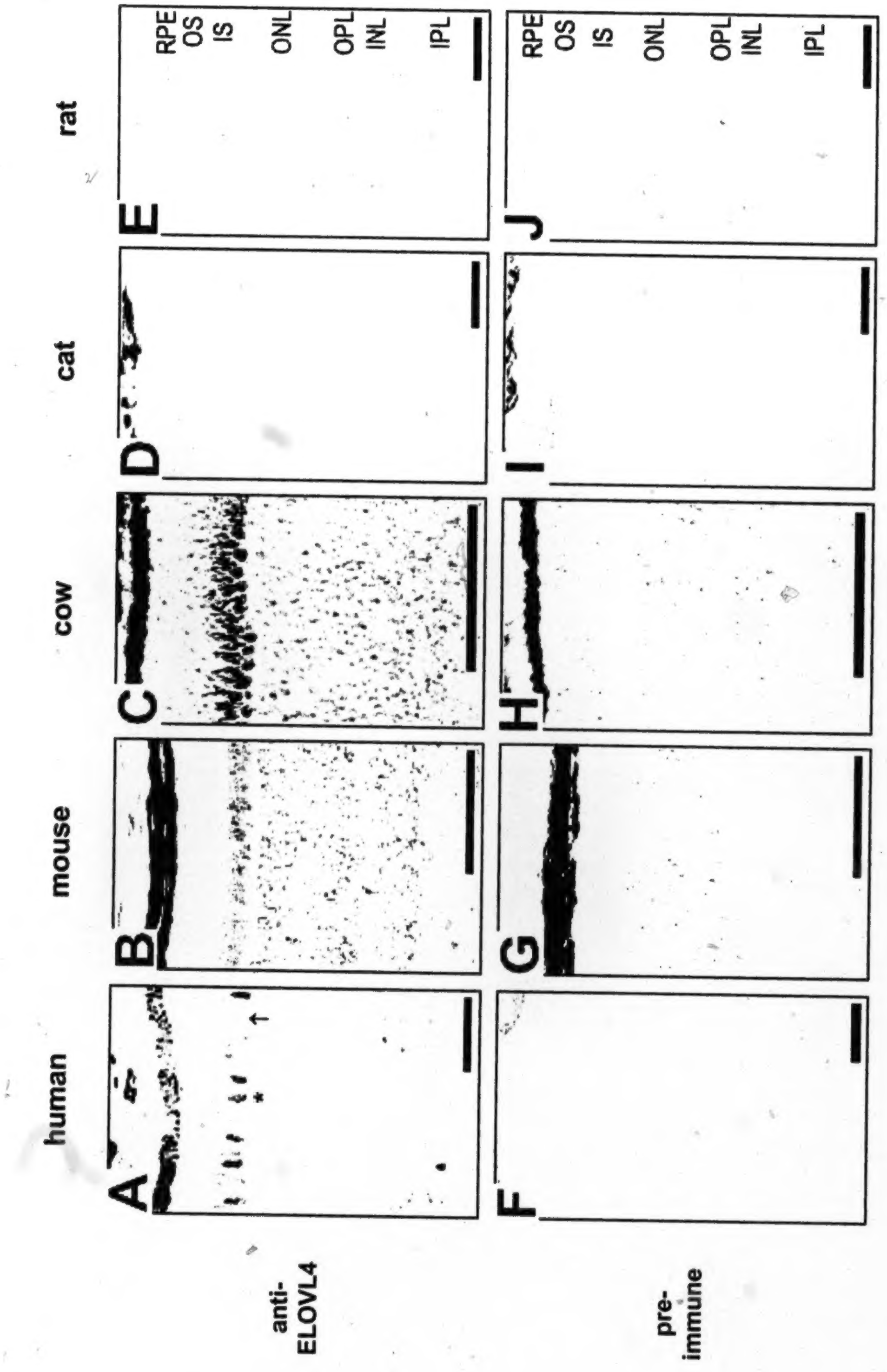

human
mouse
cow
cat
rat
- peptide
+ peptide
48 kDa
37 kDa
25 kDa
16 kDa

Figure 6.6 Conservation of sub-retinal ELOVL4 protein distribution. Immunohistochemical analysis of mammalian retina sections demonstrates conserved localization of ELOVL4 protein in the inner segments of photoreceptor cells. Retinal sections were incubated with rabbit anti-ELOVL4 polyclonal antibodies (A-E) or pre-immune serum (F-J) followed by subsequent detection and visualization with diaminobenzidine (*brown*) upon light microscopy. *A and F*, human retina sections; *B and G*, mouse retina sections; *C and H*, cow retina sections; *D and I*, cat retina sections; *E and J*, rat retina sections. Human retina sections were counterstained with methyl green to visualize cell nuclei (*blue-green*). Cone and rod photoreceptor cells are discernable in Panel A, as indicated by the asterisk and arrow, respectively. The various retinal layers are labeled as follows: RPE, retinal pigment epithelium; OS, photoreceptor outer segments; IS, photoreceptor inner segments; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer. Bars, 50 μm. Images for human sections provided by Dr. Gail M. Seigel (State University of New York at Buffalo), mouse images generated by Dr. Rajesh Ambasudhan (University of Michigan), and images for cow, cat, and rat sections provided by Dr. Jiafan Liu (University of Michigan).


human
mouse
cow
cat
rat
anti-ELOVL4
pre-immune
A
B
C
D
E
F
G
H
I
J
*
↑
RPE
OS
IS
ONL
OPL
INL
IPL

Chapter 7

Functional Analysis of the ELOVL4 Protein by Mutagenesis and Subcellular Localization Studies

Parts of this chapter have been published.
Ambasudhan R, Wang X, Jablonski MM, Thompson DA, Lagali PS, Wong PW, Sieving PA, Ayyagari R. Atrophic macular degeneration mutations in *ELOVL4* result in the intracellular misrouting of the protein. *Genomics* 2004;83(4):615-625.

7.1 INTRODUCTION

Macular degenerative diseases are a heterogeneous group of hereditary disorders that primarily affect the central region of the retina known as the macula, and consequently lead to impairments in central vision. Macular degeneration is the leading cause of severe visual loss in the developed world.[1] Several clinically distinct forms of inherited macular dystrophy have been described, and considerable variability exists among these diseases with respect to age of onset, phenotypic characteristics, and inheritance patterns.[2] Macular degenerative disorders also exhibit significant genetic heterogeneity, as multiple disease-associated mutations have been identified in seven genes thus far: *RDS*/peripherin,[3,4] *TIMP3*,[5] *ABCA4*,[6] *VMD2*/bestrophin,[7] *EFEMP1*,[8] *ELOVL4*,[9,10] and *RPGR*.[11] Tremendous progress has been made in recent years towards understanding the biological functions of these genes and their encoded proteins as well as their potential role in disease. However, based on the current state of knowledge, a common mechanism for the selective degeneration of the macula due to genetic mutations in macular dystrophy patients can not yet be deduced.

Autosomal dominant atrophic macular degeneration (adMD) with clinical features resembling autosomal recessive Stargardt disease (STGD1, OMIM #248200) is a relatively rare macular disease exhibiting a variable phenotype.[12,13] Two distinct loci have been identified for this disorder to date, STGD3/adMD on human chromosome 6q (OMIM #600110)[12-14] and STGD4 on chromosome 4p.[15] Cloning of the *ELOVL4* gene and identification of a 5-base pair deletion that underlies the disease in STGD3/adMD families mapping to 6q has previously been reported.[9] Subsequently, an independent study verified the same mutation in related patients with this disorder.[16] A second pathogenic mutation in *ELOVL4* was identified in an unrelated pedigree and is considered to be a complex mutation, consisting of two single base pair deletions positioned four base pairs apart, and is located in the same region as the 5-base pair deletion initially reported for STGD3/adMD patients.[10] Both of the known macular disease-causing mutations cause a shift in the reading frame that results in carboxy-terminal substitution and truncation of the encoded protein. The amino acids missing from the mutant proteins include those that comprise a putative endoplasmic reticulum (ER)-retention/retrieval signal defined by a dilysine motif, KXKXX (where K denotes a lysine residue and X represents any other amino acid),[17-19] that is normally found at the carboxyl terminus of the wild-type protein. Subsequent cellular mislocalization of the aberrant proteins is thought to inhibit ELOVL4 function in the macular dystrophy patients harbouring these gene mutations, which in turn leads to photoreceptor and retinal pigment epithelial cell degeneration by an as yet uncharacterized mechanism.

Although the cellular function of ELOVL4 is unknown, homology observed with members of a large conserved family of proteins[20] and the recognition of characteristic domains in the ELOVL4 amino acid sequence[9] suggest that it is an ER-bound transmembrane protein associated with the fatty acid chain elongation machinery. Retinal tissue has a unique fatty acid composition and is particularly rich in very long-chain polyunsaturated fatty acids.[21] Furthermore, it has been shown that the lipid environment is critical for the normal function of retinal proteins,[22,23] and alterations in retinal lipid composition have been observed in various retinal disorders.[24-27] Therefore it is perhaps not surprising that mutations in a predicted lipid biosynthetic gene lead to the retinal dysfunction observed in STGD3/adMD patients. ELOVL4 is the only component of the cellular fatty acid metabolism machinery that has been implicated in causing macular degeneration. It is therefore important to study the role of this protein in retinal biology as well as the functional consequences of pathogenic *ELOVL4* mutations to understand the role of fatty acids in retinal degeneration.

This report describes the development of an expression system used to analyze functional properties of wild-type and mutant versions of the ELOVL4 protein, particularly for investigation of the subcellular localization of the protein that is presumed to be critical for its putative biological function. Mammalian expression constructs encoding the ELOVL4 protein fused to a green fluorescent protein (GFP) tag[28] readily enable visualization of the expressed recombinant proteins in transiently transfected cultured cells. The effects of the known macular disease-causing mutations, as well as an engineered mutation disrupting the presumed ER-retention/retrieval signal, on intracellular distribution of ELOVL4 are examined. Additional sequence alterations introduced into the wild-type ELOVL4 cDNA provide the tools to further assess the importance of other predicted functional domains in the ELOVL4 protein. The results observed provide initial clues to understanding the molecular basis of macular degeneration associated with *ELOVL4* mutations.

7.2 MATERIALS AND METHODS

7.2.1 ELOVL4 Protein Sequence Analysis

The human ELOVL4 amino acid sequence (GenBank accession number AAH38506) was analyzed with protein sequence analysis tools accessed from the ExPASy Molecular Biology server (http://ca.expasy.org/tools/). To predict the location of potential glycosylation sites, the NetNGlyc program (version 1.0) (http://www.cbs.dtu.dk/services/NetNGlyc/), used for prediction of N-linked glycosylation sites in human proteins,[29] and the YinOYang program (version 1.2) (http://www.cbs.dtu.dk/services/YinOYang/), used for prediction of O-β-GlcNAc attachment sites in eukaryotic protein sequences,[30] were employed. The NetPhos program (version 2.0) (http://www.cbs.dtu.dk/services/NetPhos/)[31] was used to detect the presence of putative phosphorylation sites in the ELOVL4 protein. Other functional domains suggested by the primary sequence such as transmembrane segments, a histidine box motif, and a dilysine targeting signal for ER transmembrane proteins have been previously described.[9]

7.2.2 ELOVL4 cDNA Constructs and Site-Directed Mutagenesis

The wild-type cDNA sequence corresponding to the entire coding region of the human *ELOVL4* gene was amplified by PCR from a human brain cDNA library by Dr. Jiafan Liu (University of Michigan, Ann Arbor, MI) using the forward primer 5'-CTCGGATCCGCGATGGGGCTCCTGGAC-3' and reverse primer 5'-GCGGATCCCAGTTCAATTTAATCTCCTTTTGCTTTTCC-3', both containing *Bam*HI sites for insertion into the pCR II TA Cloning vector (Invitrogen Corp., Carlsbad, CA). Following sequence confirmation, the ELOVL4 cDNA was subcloned into the pEGFP-C1 mammalian expression vector (BD Biosciences Clontech) by Dr. Rajesh Ambasudhan (University of Michigan) by amplification from the pCR II-ELOVL4 plasmid using the same primers and *Pfu* DNA polymerase (Invitrogen Corp.), subsequent digestion of the product with *Bam*HI, and non-directional ligation into the *Bgl*II and *Bam*HI sites of the linearized pEGFP-C1 vector. Plasmids containing the correct orientation of the insert were verified by sequencing.

All ELOVL4 cDNA sequence variants were generated by site-directed mutagenesis using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's protocol and with mutagenic oligonucleotide primers purified by polyacrylamide gel electrophoresis. The 790-794delAACTT mutation was introduced into the wild-type cDNA sequence using forward primer 5'-ATCAGCTTCATATTTCTCTTTCTTCTACA-3' and reverse primer 5'-GGCTCTTTGTATGTCCGAATGTAGAAGAAAGAGAAATATGAAGCTGAT-3'. The 789delT+794delT mutation was created using forward primer 5'-GCCTATGCAATCAGCTTCATATTTCTCTTTCTAACTCTACATTCGGAC-3' and reverse primer 5'-GTCCGAATGTAGAGTTAGAAAGAGAAATATGAAGCTGATTGCATAGGC-3'. The 814G→C (E272Q) sequence variant was generated using forward primer 5'-TTCTACATTCGGACATACAAACAGCCTAAGAAACCAAAAGCTGG-3' and reverse primer 5'-AGCTTTTGGTTTCTTAGGCTGTTTGTATGTCCGAATGTAGAAG-3'. The 317A→T (Y106F) engineered mutation disrupting a predicted phosphorylation site in ELOVL4 was created using forward primer 5'-GGGATCATATAATGCGGGATTTAGCTATATTTGCCAGAG-3' and reverse primer 5'-CTCTGGCAAATATAGCTAAATCCCGCATTATATGATCCC-3'. The 344A→T (Y115F) putative phosphorylation site mutation was created using forward primer 5'-GCCAGAGTGTGGATTTTTCTAATAATGTTCATGAAGTCAGG-3' and reverse primer 5'-CCTGACTTCATGAACATTATTAGAAAAATCCACACTCTGGC-3'. The 472-473CA→GC+481-482CA→GC+484-485CA→GC (H158A+H161A+H162A) sequence alteration abolishing the histidine box motif was generated with forward primer 5'-CAACCAAGTTTCTTTCCTTGCTGTGTATGCTGCCTGTACGATGTTTACCTTGTGGTGG-3' and reverse primer 5'-CCACCACAAGGTAAACATCGTACAGGCAGCATACACAGCAAGGAAAGAAACTTGGTTG-3'. The 929A→G+935A→G (K310R+K312R) mutation disrupting the putative ER-retention/retrieval signal was introduced using forward primer 5'-GCAGAAAAATGGAAGAGCAAGAGGAGATTAAATTGAACTGGGCC-3' and reverse primer 5'-GGCCCAGTTCAATTTAATCTCCTCTTGCTCTTCCATTTTTCTGC-3'. For generating the del928-942 (ΔKAKGD) construct, the forward primer 5'-CCACGGGCTCCTGGACTCGG-3' and reverse primer 5'-TTATCCATTTTTCTGCTTTTTTCCATTTTC-3' were used by Dr. Rajesh Ambasudhan. Sequence integrity, in-frame cloning with the GFP fusion tag, and absence of any PCR- and cloning-introduced sequence artifacts were confirmed for all of the constructs by restriction digest analysis, PCR, and direct sequencing of the entire ELOVL4 cDNA inserts.

7.2.3 Mammalian Cell Culture and Transfection

The mutagenized pEGFP-ELOVL4 expression constructs were sent to Dr. Radha Ayyagari at the University of Michigan for subsequent transfection studies in mammalian cell cultures. African green monkey kidney (COS-7) cells were commercially obtained (American Type Culture Collection, Rockville, MD) and grown in Dulbecco's modified Eagles medium supplemented with 10% fetal bovine serum and antibiotics (GIBCO Invitrogen Corp.) in T-75 flasks. Cells were seeded on a Lab-Tek II dual-chamber slide (Nalge Nunc, Naperville, IL), grown to 80-90% confluence, and then transiently transfected with pEGFP-ELOVL4 expression constructs or the empty pEGFP-C1 vector using Lipofectamine Plus transfection reagent (Invitrogen Corp.) following the manufacturer's instructions. Forty-eight hours post-transfection, the cells were washed several times with phosphate buffered saline (PBS; pH 7.2; Sigma-Aldrich Corp.) and processed for either SDS-PAGE or immunofluorescence staining and microscopy. Cell culture work and transfections were performed by Dr. Rajesh Ambasudhan.

7.2.4 SDS-PAGE and Immunoblotting

Transfected cells were washed in ice-cold PBS and harvested into lysis buffer containing 0.25 M sucrose, 25 mM Tris-acetate (pH 7.0), and 1 mM dithiothreitol supplemented with Complete Mini Protease Inhibitor Cocktail (Roche Diagnostics Corp.). Lysates were incubated on ice for 5 minutes followed by homogenization and sonication at 4°C. Proteins were quantitated using BCA Reagent (Pierce Biotechnology). Equal amounts of protein contained in different lysates were resolved in 8% SDS-PAGE gels. Western blot analysis was carried out as previously described[20] using affinity-purified rabbit polyclonal anti-ELOVL4 antibody[20] (1:1000 dilution) or rabbit polyclonal anti-GFP antibody (1:5000 dilution; BD Biosciences Clontech) as primary antibodies and horseradish peroxidase-conjugated mouse monoclonal anti-rabbit IgG (dilution 1:5000; Pierce Biotechnology) as the secondary antibody. Chemiluminescent signals were detected using an ECL Kit (Amersham Biosciences, Piscataway, NJ) following the manufacturer's instructions. SDS-PAGE and immunoblotting procedures were performed by Dr. Rajesh Ambasudhan.

7.2.5 Immunofluorescence Labeling and Microscopy

For detection of the GFP-generated fluorescence signal alone, transfected cells were rinsed several times in cold PBS and then mounted onto slides in Vectashield anti-fade mounting medium (Vector Laboratories, Inc.) prior to observation under a Nikon-E800 epifluorescence microscope. For the double-labeling experiments, PBS-rinsed transfected cells were fixed in absolute methanol at -10°C for 15 minutes and then incubated for 30 minutes in blocking solution containing 5% normal goat serum, 1% IgG-free bovine serum albumin (BSA; Jackson ImmunoResearch Laboratories, West Grove, PA) in PBS. The cells were further incubated for 1 hour at room temperature in 4 μg/ml of mouse monoclonal anti-protein disulfide isomerase (PDI) primary antibody (StressGen Biotechnologies Corp., Victoria, BC) diluted in PBS containing 0.1% BSA. Following washes with PBS, the cells were incubated for 1 hour at room temperature in anti-mouse IgG-tetramethylrhodamine isothiocyanate (TRITC) secondary antibody (dilution 1:400; Sigma-Aldrich Corp.). Cells were again washed and then mounted in Vectashield anti-fade mounting medium (Vector Laboratories, Inc.) containing 4'6-diamidino-2-phenylindole-2HCl (DAPI) prior to microscopy. Images were captured using a Nikon DMX 1200 digital camera and Nikon ACT-1 image collection software. Immunoflurescence experiments were carried out by Dr. Rajesh Ambasudhan.

7.3 RESULTS

7.3.1 Mutagenesis of Putative Functional Domains in ELOVL4

Bioinformatic amino acid sequence analysis of the ELOVL4 protein revealed the presence of a number of predicted functional domains, including O-linked glycosylation sites at positions 16, 229, and 292, an N-linked glycosylation site at residue 20, six phosphorylation sites at amino acid positions 106, 115, 139, 269, 292, and 294, and previously described histidine box motif encompassing residues 158-162 and a dilysine ER-retention/retrieval signal at carboxy-terminal residues 310-314 (Figure 7.1). The two autosomal dominant macular disease-causing mutations identified in the *ELOVL4* gene to date cause a frameshift that terminates the wild-type protein sequence at residue 263. Other DNA sequence alterations in the *ELOVL4* gene have been described among some age-related macular degeneration (AMD) patients, including a variant that encodes a missense glutamic acid to glutamine amino acid substitution at position 272.[32] To begin to address the importance of all these sequence elements and their contributions to *ELOVL4* function, site-directed mutagenesis was performed to disrupt selected putative functional domains in the wild-type protein and to introduce the disease-associated sequence variations into the ELOVL4 cDNA. Sequencing of the mutagenesis reaction products and comparison with the wild-type ELOVL4 coding region template confirmed the presence of the correct nucleotide replacements or deletions in the altered cDNAs (Figures 7.2 and 7.3).

7.3.2 Expression of GFP-ELOVL4 Fusion Proteins in COS-7 Cells

For the purposes of transfecting the wild-type and mutagenized cDNAs into cultured cells, they were inserted into the pEGFP-C1 mammalian expression vector (Figure 7.4A). This vector contains the coding sequence for a GFP tag that is linked to the amino terminus of the ELOVL4 proteins upon translation in mammalian cells. Thus expression of the fusion protein and detection of the fluorescent signal enables an assessment of the subcellular localization of ELOVL4. To determine the influence of the pathogenic deletions identified in STGD3/adMD patients on the intracellular trafficking of the ELOVL4 protein, the wild-type, macular disease mutation-containing (790-794delAACTT; 789delT+794delT) and ER-targeting signal-disrupted (del928-942/ΔKAKGD) versions of the ELOVL4 cDNA constructs (Figure 7.4B) were transiently transfected into COS-7 cells.

Following transfection, cell lysates were prepared and the constituent proteins were analyzed to verify synthesis of the correct GFP-ELOVL4 fusion proteins in appropriately transfected cells (Figure 7.5). Immunoblot analysis of the extracts from COS-7 cells expressing the GFP-wild-type ELOVL4 fusion protein revealed an approximately 67-kilodalton band (Figures 7.5A and 7.5B, lane 3) that corresponds to the expected size for the combined ELOVL4 (~37 kDa, Figure 7.5A, lane 1) and GFP (~31 kDa, Figure 7.5B, lane 2) proteins. A band at 63 kDa was observed for extracts from cells expressing each of the macular disease-causing mutations in ELOVL4

(Figures 7.5A and 7.5B, lanes 4 and 5), corresponding to the expected molecular mass of the truncated protein (ELOVL4 790-794delAACTT/789delT+794delT, ~32.1 kDa + GFP, ~31 kDa). A band of expected size was also observed in cells transfected with constructs encoding ELOVL4 lacking the carboxy-terminal ER-retention/retrieval signal (Figures 7.5A and 7.5B, lane 6). The expected bands corresponding to each of the fusion proteins were detected using both anti-ELOVL4 (Figure 7.5A) and anti-GFP (Figure 7.5B) antibodies.

7.3.3 *Intracellular Distribution of Wild-type ELOVL4 in COS-7 Cells*

To determine the subcellular localization of the wild-type ELOVL4 protein, immunofluorescence microscopy of the transiently transfected COS-7 cells was performed (Figure 7.6). Cells expressing the GFP-ELOVL4 fusion protein were examined for the intrinsic GFP fluorescence signal. Micrographs of the transfected cells revealed that the GFP-wild-type ELOVL4 fusion product was distributed in a distinct perinuclear and reticular pattern (Figure 7.6A). This staining pattern is consistent with ER localization as indicated by the similar results obtained when the cells were immunolabeled with antibodies recognizing the ER marker protein, protein disulfide isomerase (PDI)[33] (Figure 7.6B). The GFP-mediated fluorescence and PDI-labeling patterns were virtually indistinguishable in these cells, as verified by the near complete overlap observed upon overlaying the two images (Figure 7.6C). This confirms the ER localization of the wild-type ELOVL4 protein. As expected, control experiments on mock-transfected cells revealed no signal for ELOVL4 (Figure 7.6D), and cells transfected with the empty pEGFP-C1 vector exhibited a diffuse staining pattern (Figure 7.6G) markedly different from the fluorescence signal generated by the GFP-ELOVL4 fusion protein. Furthermore, no significant co-localization of the ELOVL4 protein with the DAPI-stained nucleus was observed (Figure 7.6C). These results indicate that wild-type ELOVL4 is predominantly a resident of the ER in mammalian cells.

7.3.4 *Subcellular Localization of Mutant ELOVL4 Proteins Carrying Macular Disease-Causing Deletions*

The COS-7 cells expressing the 5-base pair deletion or the complex mutation in the coding region of the ELOVL4 gene were similarly analyzed by immunofluorescence microscopy. Unlike the wild-type protein, both of the proteins harbouring macular disease mutations showed strong staining in a very localized region near the nucleus (Figures 7.6J, 7.6M) in a vast majority of the cells. Double-labeling revealed that these mutant proteins do not appreciably co-localize with PDI (Figures 7.6L and 7.6O). Thus it appears that the pathogenic mutations in ELOVL4 result in altered subcellular distribution of the protein, disrupting its normal localization to the ER.

7.3.5 ER Localization of ELOVL4 Depends on the Carboxy-Terminal Dilysine Motif

It is predicted that both of the macular disease-causing mutations identified in the ELOVL4 gene lead to loss of the carboxy-terminal dilysine ER-targeting motif in the encoded protein and that this may underlie their pathogenicity. To test whether the wild-type ELOVL4 localization observed is indeed dependent on the putative ER-retention/retrieval signal and if its absence results in a similar intracellular distribution as the mutant proteins, the GFP-tagged ELOVL4 cDNA construct carrying the targeted deletion of the dilysine motif was transiently expressed in COS-7 cells. Immunofluorescence analysis of the transfected cells revealed localized GFP fluorescence predominantly in juxtanuclear regions (Figure 7.6P). Double-labeling of these cells with the anti-PDI antibody (Figure 7.6R) demonstrated a distinctly different distribution pattern than that observed for the ER marker protein alone (Figure 7.6Q) and was reminiscent of the localized staining detected for the macular disease mutations (Figures 7.6L and 7.6O). These findings indicate that the KAKGD motif at the carboxyl terminus of ELOVL4 is involved in the retention/retrieval of the protein to the ER, and provides further evidence that ELOVL4 mutants lacking this motif are subject to intracellular misrouting.

7.4 DISCUSSION

While mutations in the *ELOVL4* gene are known to underlie different forms of macular degeneration, the mechanism by which these mutations disrupt *ELOVL4* function and ultimately cause vision loss has only been speculative. The studies described here provide insight into STGD3/adMD disease etiology through three important observations: localization of the wild-type ELOVL4 protein to the ER, the dependence of ER localization of ELOVL4 on the presence of the carboxy-terminal dilysine motif, KAKGD, and defective intracellular trafficking of mutant ELOVL4 proteins associated with STGD3/adMD patients. These findings confirm the prediction of aberrant mutant ELOVL4 protein subcellular localization that was previously made based on the observation that the two mutations found in macular dystrophy patients could lead to carboxy-terminal truncation and consequent loss of the putative ER-targeting signal.

Signal-mediated sorting of proteins has been implicated as an important means of ER quality control for protein trafficking.[34] The presence of a dilysine motif located at either -3 and -4 (KKXX) or at -3 and -5 (KXKXX) positions from the carboxyl terminus has been reported as an ER-retention/retrieval signal for many ER membrane proteins, for which it functions in COP-1-mediated retrograde transport from post-ER compartments.[17-19] However, membrane proteins carrying an intact dilysine motif are also found to be localized to regions other than the ER,[35-40] indicating that the mere presence of the above motif does not ensure localization of a protein to the ER. The observation of the absence of ER localization of ELOVL4 proteins lacking the dilysine signal suggests the functionality of the carboxy-terminal KAKGD motif in this protein as an ER-retention/retrieval signal.

Misrouting of the mutant ELOVL4 proteins implies that they are destined for a subcellular compartment other than the ER. Additional experiments performed on the COS-7 cells transiently transfected with the GFP-ELOVL4 expression constructs described here indicate that the mutant proteins accumulate in the Golgi apparatus, and more specifically in the *trans* Golgi network (TGN).[41] Several studies have described the export of ER proteins to the Golgi from where they are subsequently retrieved.[35,42] It is possible that such a mechanism may also operate for ELOVL4. In support of this hypothesis, immunofluorescence staining has shown a minor zone of overlap between wild-type ELOVL4 and Golgi marker proteins,[41] and bioinformatic analysis of the ELOVL4 protein sequence predicted the presence of three potential sites for O-linked glycosylation, a protein modification that takes place in the Golgi apparatus.[43] This would suggest the possibility that the normal cellular routing of the ELOVL4 protein could involve transport to the Golgi and signal-mediated retrieval back to the ER. In a selective or receptor-mediated transport model of intracellular protein trafficking, the absence of the ER-retrieval signal and other signals leading to export from the Golgi may result in defective accumulation of the mutant ELOVL4 proteins within the TGN, the site of protein sorting for subsequent transport.[44] Alternatively, Golgi localization may be mediated by the creation of a Golgi-localization signal within the mutant proteins or through their aggregation with resident components of this subcellular compartment via novel protein-protein or protein-lipid interactions. The transmembrane domains of many Golgi proteins, such as glycosyltransferases, act as dominant localization signals independent of specific motifs commonly occurring within the amino acid sequences comprising these domains.[44-46] The disease-causing mutations that affect the fifth putative transmembrane segment in ELOVL4 may generate alterations in the protein that similarly target them to the Golgi apparatus. An aggregation or oligomerization model has also been proposed to play a role in the localization of glycosyltransferases and other Golgi-resident proteins that suggests the existence of interactions between transmembrane domains of the proteins and the lipid bilayers of Golgi membranes.[47,48] It has also been suggested that large hetero-oligomeric structures are formed via interactions between different proteins in the same Golgi compartment that result in their exclusion from forward moving vesicles and thus further anterograde transport through the secretory pathway.[49] These models may also account for the mislocalization of the mutant ELOVL4 proteins in which a significant portion of the carboxy-terminal tail is lost and thus may sufficiently alter the conformation to allow for otherwise absent intermolecular interactions occurring in the Golgi complex. Further understanding the specific mechanism by which the aberrant subcellular localization arises as well as the fate of the Golgi-accumulated mutant proteins may help determine the role of ELOVL4 mutations in causing degeneration of the macula.

Defective intracellular trafficking of proteins has been reported to be associated with several human diseases, including those involving degeneration of neuronal tissue and/or the visual system.[50-53] STGD3/adMD could be a new addition to this fast-growing list of disorders. It is interesting to note that some of the rhodopsin mutations associated with autosomal dominant retinitis pigmentosa also result in defective protein trafficking, and that some of the mutant proteins have been shown to accumulate in the ER or Golgi or form aggresomes.[54-56] However, it is unclear as to how these cellular changes lead to photoreceptor degeneration. Although ELOVL4 is expressed in both rod and cone photoreceptor cells,[9,20] no mutations have thus far been found in this gene in patients with peripheral retinal disease.[57,58] It is intriguing to note that altered trafficking of the photoreceptor-expressed proteins rhodopsin and ELOVL4 results in selective degeneration of the peripheral retina or the macula, respectively. Discerning the role of these gene products in various cellular pathways would provide further insights into the mechanism behind this phenomenon.

The ELOVL4 protein has been predicted to play a role in the elongation of very long-chain fatty acids based on its homology with a large family of other proteins whose members have been shown to be involved in fatty acid elongation.[20] In animal cells, synthesis of very long-chain fatty acids occurs through a series of sequential desaturation and chain-elongation events using shorter-chained substrates.[59] Chain-elongation steps take place in the membrane of the ER,[59] to which ELOVL4 is localized. Although the data presented in this study do not provide direct evidence to determine its biological function, they further support the prediction that ELOVL4 participates in the elongation of long-chain fatty acids based on the known subcellular location of this metabolic process.

Aside from the ER-targeting signal, a number of other potentially functional domains have been identified in ELOVL4 that may contribute to protein structure and/or activity. Three O-linked glycosylation sites and one N-linked glycosylation site are predicted to lie within the protein sequence, which may be involved in protein folding and stability, membrane associations, substrate recognition and affinity, or catalytic properties. The well-studied retinal proteins rhodopsin and ABCR are known to be glycosylated,[60,61] and mutations affecting consensus glycosylation sites have been identified in the genes encoding each of these proteins in retinal dystrophy patients.[62,63] Furthermore, intraocular injection of tunicamycin, an inhibitor of formation of the asparagine-linked oligosaccharides of glycoproteins, has been demonstrated to induce photoreceptor-specific degeneration in frogs,[64] underscoring the importance of this type of protein modification on retinal cell viability. Phosphorylation is another common post-translational modification of proteins and can play a variety of functional roles. In addition to structural effects and possible regulatory functions, phosphorylation at residues near the carboxyl terminus has also been implicated in retrograde transport of membrane proteins from the Golgi to the ER,[65] a transport process presumed to occur for ELOVL4 based on our observations. The carboxy-terminal truncation of the ELOVL4 protein due to each of the macular disease-causing mutations results in the loss of three out of six of the predicted phosphorylation sites, and these amino acids may therefore also be important for proper intracellular trafficking of ELOVL4. The HXXHH histidine box motif is a defining feature and critical determinant of enzymatic activity in various fatty acid biosynthetic proteins.[66] Mutagenesis of constituent amino acids within this motif as well as conserved flanking residues have been shown to drastically reduce enzyme activity and therefore also fatty acid product yield.[66,67] Assessment of each of these putative functional sequence elements in the ELOVL4 protein will enhance our understanding of its mode of action in retinal cells and may reveal possible disease mechanisms for mutations that may be associated with these protein domains in additional retinal dystrophy patients.

The 814G→C *ELOVL4* gene sequence variant found in some AMD patients, encoding a E272Q amino acid substitution, may also impart functional consequences on the ELOVL4 protein. The replacement of a negatively charged glutamic acid residue with an uncharged glutamine residue might have significant effects on both intramolecular and intermolecular interactions. E→Q missense mutations have been observed in other cases of neurodegenerative disease[68,69] and in some cases have been implicated in the synthesis of aberrant proteins.[70,71] Deducing the effects of this sequence variation as well as others that have been reported to occur in cases of AMD may uncover additional functional regions of the protein that are not obviously suggested by the amino acid sequence.

Long-chained polyunsaturated fatty acids have been implicated to play a significant biological role in maintaining the normal function of the retina. Although many retinal disease genes have been cloned to date, none of them have been associated with fatty acid biosynthesis. Unraveling the role of *ELOVL4* gene mutations in causing macular degeneration may provide a direct link between fatty acid metabolism and photoreceptor cell loss. Determining the biological function of ELOVL4, which depends critically on its intracellular location and undoubtedly other characteristics defined by its amino acid sequence, would thus not only provide further clues to pathogenetic mechanisms occurring in STGD3/adMD patients, but would also expand our basic knowledge of retinal cell biology and may suggest avenues for therapeutic intervention in these and other forms of retinal disease.

7.5 ACKNOWLEDGEMENTS

The authors thank Jiafan Liu and Laura E. Kakuk (University of Michigan) for their participation during the initial phase of the experiments and David Allen (University of Michigan) for technical advice. This work was supported by grants from the NIH

(EY13198), Foundation Fighting Blindness, USA, Research to Prevent Blindness, Inc., USA (RA); Canadian Institutes of Health Research (PSL); E.A. Baker Foundation of the Canadian National Institute for the Blind, Alberta Heritage Foundation for Medical Research (PSL, PWW); Foundation Fighting Blindness, Canada (PWW); and NIH core grants to the University of Michigan, Department of Ophthalmology (EY07003), and Vision Research (EY07060).

7.6 REFERENCES

1. van Leeuwen R, Klaver CC, Vingerling JR, Hofman A, de Jong PT. (2003) Epidemiology of age-related maculopathy: a review. *Eur. J. Epidemiol.* 18:845-854.

2. Daiger SP. RetNet, the Retinal Information Network http://www.sph.uth.tmc.edu/Retnet/ .

3. Wells J, Wroblewski J, Keen J, Inglehearn C, Jubb C, Eckstein A, Jay M, Arden G, Bhattacharya S, Fitzke F, Bird A. (1993) Mutations in the human retinal degeneration slow (*RDS*) gene can cause either retinitis pigmentosa or macular dystrophy. *Nat. Genet.* 3:213-218.

4. Kohl S, Christ-Adler M, Apfelstedt-Sylla E, Kellner U, Eckstein A, Zrenner E, Wissinger B. (1997) *RDS/*peripherin gene mutations are frequent causes of central retinal dystrophies. *J. Med. Genet.* 34:620-626.

5. Weber BH, Vogt G, Pruett RC, Stohr H, Felbor U. (1994) Mutations in the tissue inhibitor of metalloproteinases-3 (TIMP3) in patients with Sorsby's fundus dystrophy. *Nat. Genet.* 8:352-356.

6. Allikmets R, Singh N, Sun H, Shroyer NF, Hutchinson A, Chidambaram A, Gerrard B, Baird L, Stauffer D, Peiffer A, Rattner A, Smallwood P, Li Y, Anderson KL, Lewis RA, Nathans J, Leppert M, Dean M, Lupski JR. (1997) A photoreceptor cell-specific ATP-binding transporter gene (*ABCR*) is mutated in recessive Stargardt macular dystrophy. *Nat. Genet.* 15:236-246.

7. Petrukhin K, Koisti MJ, Bakall B, Li W, Xie G, Marknell T, Sandgren O, Forsman K, Holmgren G, Andreasson S, Vujic M, Bergen AA, McGarty-Dugan V, Figueroa D, Austin CP, Metzker ML, Caskey CT, Wadelius C. (1998) Identification of the gene responsible for Best macular dystrophy. *Nat. Genet.* 19:241-247.

8. Stone EM, Lotery AJ, Munier FL, Heon E, Piguet B, Guymer RH, Vandenburgh K, Cousin P, Nishimura D, Swiderski RE, Silvestri G, Mackey DA, Hageman GS, Bird AC, Sheffield VC, Schorderet DF. (1999) A single *EFEMP1* mutation associated with both Malattia Leventinese and Doyne honeycomb retinal dystrophy. *Nat. Genet.* 22:199-202.

9. Zhang K, Kniazeva M, Han M, Li W, Yu Z, Yang Z, Li Y, Metzker ML, Allikmets R, Zack DJ, Kakuk LE, Lagali PS, Wong PW, MacDonald IM, Sieving PA, Figueroa DJ, Austin CP, Gould RJ, Ayyagari R, Petrukhin K. (2001) A 5-bp deletion in *ELOVL4* is associated with two related forms of autosomal dominant macular dystrophy. *Nature Genet.* 27:89-93.

10. Bernstein PS, Tammur J, Singh N, Hutchinson A, Dixon M, Pappas CM, Zabriskie NA, Zhang K, Petrukhin K, Leppert M, Allikmets R. (2001) Diverse macular dystrophy phenotype caused by a novel complex mutation in the *ELOVL4* gene. *Invest. Ophthalmol. Vis. Sci.* 42:3331-3336.

11. Ayyagari R, Demirci FY, Liu J, Bingham EL, Stringham H, Kakuk LE, Boehnke M, Gorin MB, Richards JE, Sieving PA. (2002) X-linked recessive atrophic macular degeneration from *RPGR* mutation. *Genomics* 80:166-171.

12. Griesinger IB, Sieving PA, Ayyagari R. (2000) Autosomal dominant macular atrophy at 6q14 excludes CORD7 and MCDR1/PBCRA loci. *Invest. Ophthalmol. Vis. Sci.* 41:248-255.

13. Lagali PS, MacDonald IM, Griesinger IB, Chambers ML, Ayyagari R, Wong PW. (2000) Autosomal dominant Stargardt-like macular dystrophy segregating in a large Canadian family. *Can. J. Ophthalmol.* 35:315-24.

14. Stone EM, Nichols BE, Kimura AE, Weingeist TA, Drack A, Sheffield VC. (1994) Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. *Arch. Ophthalmol.* 112:765-772.

15. Kniazeva M, Chiang MF, Morgan B, Anduze AL, Zack DJ, Han M, Zhang K. (1999) A new locus for autosomal dominant stargardt-like disease maps to chromosome 4. *Am. J. Hum. Genet.* 64:1394-1399.

16. Edwards AO, Donoso LA, Ritter R 3rd. (2001) A novel gene for autosomal dominant Stargardt-like macular dystrophy with homology to the SUR4 protein family. *Invest. Ophthalmol. Vis. Sci.* 42:2652-2663.

17. Jackson MR, Nilsson T, Peterson PA. (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO J.* 9:3153-3162.

18. Jackson MR, Nilsson T, Peterson PA. (1993) Retrieval of transmembrane proteins to the endoplasmic reticulum. *J. Cell. Biol.* 121:317-333.

19. Teasdale RD, Jackson MR. (1993) Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the golgi apparatus. *Annu. Rev. Cell. Dev. Biol.* 12:27-54.

20. Lagali PS, Liu J, Ambasudhan R, Kakuk LE, Bernstein SL, Seigel GM, Wong PW, Ayyagari R. (2003) Evolutionarily Conserved *ELOVL4* Gene Expression in the Vertebrate Retina. *Invest. Ophthalmol. Vis. Sci.* 44:2841-2850.

21. Fliesler SJ, Anderson RE. (1983) Chemistry and metabolism of lipids in the vertebrate retina. *Prog. Lipid. Res.* 22:79-131.

22. Brown MF. (1994) Modulation of rhodopsin function by properties of the membrane bilayer. *Chem. Phys. Lipids.* 73:159-180.

23. Ahn J, Wong JT, Molday RS. (2000) The effect of lipid environment and retinoids on the ATPase activity of ABCR, the photoreceptor ABC transporter responsible for Stargardt macular dystrophy. *J. Biol. Chem.* 275:20399-20405.

24. Gong J, Rosner B, Rees DG, Berson EL, Weigel-DiFranco CA, Schaefer EJ. (1992) Plasma docosahexaenoic acid levels in various genetic forms of retinitis pigmentosa. *Invest. Ophthalmol. Vis. Sci.* 33:2596-2602.

25. Holman RT, Bibus DM, Jeffrey GH, Smethurst P, Crofts JW. (1994) Abnormal plasma lipids of patients with Retinitis pigmentosa. *Lipids* 29:61-65.

26. Maude MB, Anderson EO, Anderson RE. (1998) Polyunsaturated fatty acids are lower in blood lipids of Usher's type I but not Usher's type II. *Invest. Ophthalmol. Vis. Sci.* 39:2164-2166.

27. Anderson RE, Maude MB, Alvarez RA, Acland G, Aguirre GD. (1999) A hypothesis to explain the reduced blood levels of docosahexaenoic acid in inherited retinal degenerations caused by mutations in genes encoding retina-specific proteins. *Lipids* 34:S235-S237.

28. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. (1994) Green fluorescent protein as a marker for gene expression. *Science* 263:802-805.

29. Gupta R, Jung E, Brunak S. Prediction of N-glycosylation sites in human proteins. (manuscript in preparation)

30. Gupta R, Hansen J, Brunak S. Identifying intracellular O-(beta)-GlcNAc `yin-yang' switches in the available human proteome. (manuscript in preparation)

31. Blom N, Gammeltoft S, Brunak S. (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* 294:1351-1362.

32. Ayyagari R, Zhang K, Hutchinson A, Yu Z, Swaroop A, Kakuk LE, Seddon JM, Bernstein PS, Lewis RA, Tammur J, Yang Z, Li Y, Zhang H, Yashar BM, Liu J, Petrukhin K, Sieving PA, Allikmets R. (2001) Evaluation of the ELOVL4 gene in patients with age-related macular degeneration. *Ophthalmic Genet.* 22:233-239.

33. Koch GL. (1987) Reticuloplasmins: a novel group of proteins in the endoplasmic reticulum. *J. Cell Sci.* 87:491-492.

34. Ellgaard L, Helenius A. (2003) Quality control in the endoplasmic reticulum. *Nat. Rev. Mol. Cell. Biol.* 4:181-191.

35. Schroder S, Schimmoller F, Singer-Kruger B, Riezman H. (1995) The Golgi-localization of yeast Emp47p depends on its di-lysine motif but is not affected by the ret1-1 mutation in alpha-COP. *J. Cell. Biol.* 131:895-912.

36. Itin C, Schindler R, Hauri HP. (1995) Targeting of protein ERGIC-53 to the ER/ERGIC/cis-Golgi recycling pathway. *J. Cell Biol.* 131:57-67.

37. Letourneur F, Hennecke S, Demolliere C, Cosson P. (1995) Steric masking of a dilysine endoplasmic reticulum retention motif during assembly of the human high affinity receptor for immunoglobulin E. *J. Cell Biol.* 129:971-978.

38. Presley JF, Cole NB, Schroer TA, Hirschberg K, Zaal KJ, Lippincott-Schwartz J. (1997) ER-to-Golgi transport visualized in living cells. *Nature* 389:81-85.

39. Lowe SL, Wong SH, Hong W. (1996) The mammalian ARF-like protein 1 (Arl1) is associated with the Golgi complex. *J. Cell Sci.* 109:209-220.

40. Griffith KJ, Chan EK, Lung CC, Hamel JC, Guo X, Miyachi K, Fritzler MJ. (1997) Molecular cloning of a novel 97-kd Golgi complex autoantigen associated with Sjogren's syndrome. *Arthritis Rheum.* 40:1693-1702.

41. Ambasudhan R, Wang X, Jablonski MM, Thompson DA, Lagali PS, Wong PW, Sieving PA, Ayyagari R. (2004) Atrophic macular degeneration mutations in *ELOVL4* result in the intracellular misrouting of the protein. *Genomics* 83:615-625.

42. Vashist S, Kim W, Belden WJ, Spear ED, Barlowe C, Ng DT. (2001) Distinct retrieval and retention mechanisms are required for the quality control of endoplasmic reticulum protein folding. *J. Cell Biol.* 155:355-368.

43. Rottger S, White J, Wandall HH, Olivo JC, Stark A, Bennett EP, Whitehouse C, Berger EG, Clausen H, Nilsson T. (1998) Localization of three human polypeptide GalNAc-transferases in HeLa cells suggests initiation of O-linked glycosylation throughout the Golgi apparatus. *J. Cell Sci.* 111:45-60.

44. van Vliet C, Thomas EC, Merino-Trigo A, Teasdale RD, Gleeson PA. (2003) Intracellular sorting and transport of proteins. *Prog. Biophys. Mol. Biol.* 83:1-45.

45. Munro S. (1991) Sequences within and adjacent to the transmembrane segment of alpha-2,6-sialyltransferase specify Golgi retention. *EMBO J.* 10:3577-3588.

46. Masibay AS, Balaji PV, Boeggeman EE, Qasba PK. (1993) Mutational analysis of the Golgi retention signal of bovine beta-1,4-galactosyltransferase. *J. Biol. Chem.* 268:9908-9916.

47. Machamer CE. (1991) Golgi retention signals: do membranes hold the key? *Trends Cell. Biol.* 1:141-144.

48. Weisz OA, Swift AM, Machamer CE. (1993) Oligomerization of a membrane protein correlates with its retention in the Golgi complex. *J. Cell Biol.* 122:1185-1196.

49. Nilsson T, Slusarewicz P, Hoe MH, Warren G. (1993) Kin recognition. A model for the retention of Golgi enzymes. *FEBS Lett.* 330:1-4.

50. Aridor M, Hannan LA. (2002) Traffic jams II: an update of diseases of intracellular transport. *Traffic* 3:781-790.

51. Katayama T, Imaizumi K, Sato N, Miyoshi K, Kudo T, Hitomi J, Morihara T, Yoneda T, Gomi F, Mori Y, Nakano Y, Takeda J, Tsuda T, Itoyama Y, Murayama O, Takashima A, St George-Hyslop P, Takeda M, Tohyama M. (1999) Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. *Nat. Cell Biol.* 1:479-485.

52. Bauer MF, Neupert W. (2001) Import of proteins into mitochondria: a novel pathomechanism for progressive neurodegeneration. *J. Inherit. Metab. Dis.* 24:166-180.

53. Schwahn U, Paland N, Techritz S, Lenzner S, Berger W. (2001) Mutations in the X-linked RP2 gene cause intracellular misrouting and loss of the protein. *Hum. Mol. Genet.* 10:1177-1183.

54. Kaushal S, Khorana HG. (1994) Structure and function in rhodopsin. 7. Point mutations associated with autosomal dominant retinitis pigmentosa. Biochemistry 33:6121-6128.

55. Sung CH, Davenport CM, Nathans J. (1993) Rhodopsin mutations responsible for autosomal dominant retinitis pigmentosa: clustering of functional classes along the polypeptide chain. *J. Biol. Chem.* 268:26645-26649.

56. Saliba RS, Munro PM, Luthert PJ, Cheetham ME. (2002) The cellular fate of mutant rhodopsin: quality control, degradation and aggresome formation. *J. Cell Sci.* 115:2907-2918.

57. Rivolta C, Ayyagari R, Sieving PA, Berson EL, Dryja TP. (2003) Evaluation of the *ELOVL4* gene in patients with autosomal recessive retinitis pigmentosa and Leber congenital amaurosis. *Mol. Vis.* 9:49-51.

58. Li Y, Marcos I, Borrego S, Yu Z, Zhang K, Antiñolo G. (2001) Evaluation of the *ELOVL4* gene in families with retinitis pigmentosa linked to the *RP25* locus. *J. Med. Genet.* 38:478-480.

59. Cinti DL, Cook L, Nagi MN, Suneja SK. (1992) The fatty acid chain elongation system of mammalian endoplasmic reticulum. *Prog. Lipid Res.* 31:1-51.

60. Kean EL. (1999) The dolichol pathway in the retina and its involvement in the glycosylation of rhodopsin. *Biochim. Biophys. Acta* 1473:272-285.

61. Bungert S, Molday LL, Molday RS. (2001) Membrane topology of the ATP binding cassette transporter ABCR and its relationship to ABC1 and related ABCA transporters: identification of N-linked glycosylation sites. *J. Biol. Chem.* 276:23539-23546.

62. Sullivan LJ, Makris GS, Dickinson P, Mulhall LE, Forrest S, Cotton RG, Loughnan MS. (1993) A new codon 15 rhodopsin gene mutation in autosomal dominant retinitis pigmentosa is associated with sectorial disease. *Arch. Ophthalmol.* 111:1512-1517.

63. Briggs CE, Rucinski D, Rosenfeld PJ, Hirose T, Berson EL, Dryja TP. (2001) Mutations in ABCR (*ABCA4*) in patients with Stargard macular degeneration or cone-rod degeneration. *Invest. Ophthalmol. Vis. Sci.* 42:2229-2236.

64. Fliesler SJ, Rapp LM, Hollyfield JG. (1984) Photoreceptor-specific degeneration caused by tunicamycin. *Nature* 311:575-577.

65. Cabrera M, Muniz M, Hidalgo J, Vega L, Martin ME, Velasco A. (2003) The retrieval function of the KDEL receptor requires PKA phosphorylation of its C-terminus. *Mol. Biol. Cell* 14:4114-4125.

66. Shanklin J, Whittle E, Fox BG. (1994) Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* 33:12787-12794.

67. Sayanova O, Beaudoin F, Libisch B, Castel A, Shewry PR, Napier JA. (2001) Mutagenesis and heterologous expression in yeast of a plant Δ^6-fatty acid desaturase. *J. Exp. Bot.* 52:1581-1585.

68. Levy E, Carman MD, Fernandez-Madrid IJ, Power MD, Lieberburg I, van Duinen SG, Bots GT, Luyendijk W, Frangione B. (1990) Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. Science 248:1124-1126.

69. Peoc'h K, Manivet P, Beaudry P, Attane F, Besson G, Hannequin D, Delasnerie-Laupretre N, Laplanche JL. (2000) Identification of three novel mutations (E196K, V203I, E211Q) in the prion protein gene (*PRNP*) in inherited prion diseases with Creutzfeldt-Jakob disease phenotype. *Hum. Mutat.* 15:482.

70. Prelli F, Levy E, van Duinen SG, Bots GT, Luyendijk W, Frangione B. (1990) Expression of a normal and variant Alzheimer's beta-protein gene in amyloid of hereditary cerebral hemorrhage, Dutch type: DNA and protein diagnostic assays. *Biochem. Biophys. Res. Commun.* 170:301-307.

71. Watson DJ, Selkoe DJ, Teplow DB. (1999) Effects of the amyloid precursor protein Glu693-->Gln 'Dutch' mutation on the production and stability of amyloid beta-protein. *Biochem. J.* 340:703-709.

Figure 7.1 A schematic diagram depicting the putative functional protein domains in human ELOVL4 as predicted from the amino acid sequence. Amino acid positions are indicated above the diagram and are numbered from the amino-terminus (N) to the carboxyl-terminus (C). Identities of the putative domains and disease-associated sequence alterations under study are given below the diagram. Brown regions denote putative O-linked glycosylation sites (O), pink region corresponds to a potential N-linked glycosylation site (NXT), blue regions depict predicted transmembrane domains, red bars represent potential phosphorylation sites (P), gold region corresponds to a histidine box motif (HXXHH), and green region represents a potential ER-retention/retrieval signal (KXKXX). The location of the two known sequence alterations causing macular dystrophy are indicated by Δ1 (797-801delAACTT) and Δ2 (790delT+794delT), while the 821G→C (E272Q) sequence variant found among some AMD patients is designated V.

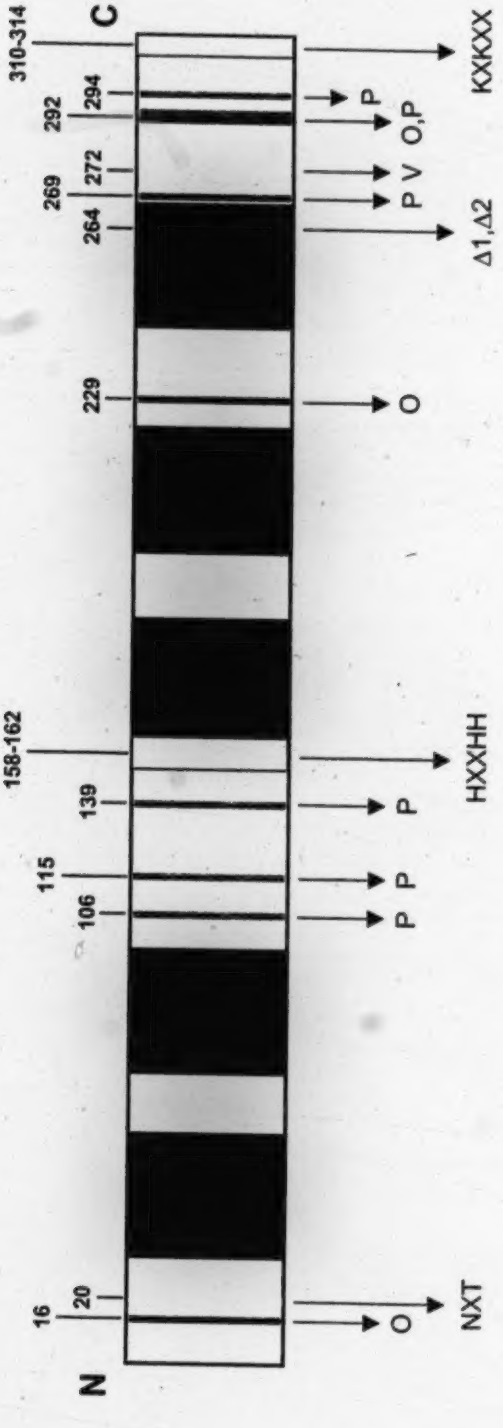
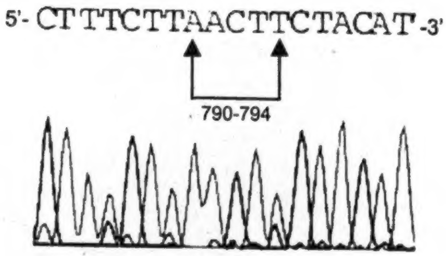
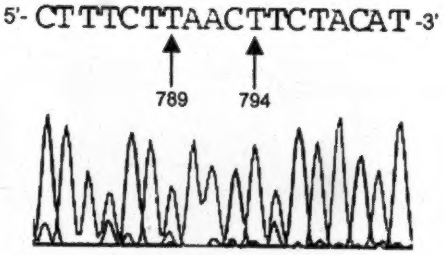

N
C
16
20
O
NXT
106
115
139
158-162
P
P
P
HXXHH
229
O
264
269
272
292
294
310-314
Δ1,Δ2
P
V
O,P
P
KXKXX

Figure 7.2 Sequence alterations in the *ELOVL4* coding sequence associated with human macular disease introduced into wild-type ELOVL4 cDNA by site-directed mutagenesis. Electrophoretograms of portions of the wild-type cDNA sequence are given on the left and those of the engineered mutations in the corresponding locations are shown on the right. Numbers refer to nucleotide positions within the ELOVL4 cDNA sequence relative to the translation start site (amino acid position within the encoded protein is given in parentheses).

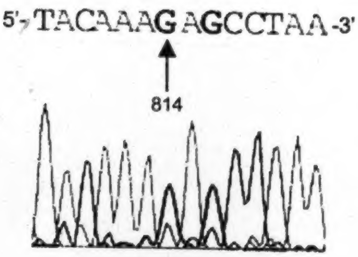
Wild-type ELOVL4 cDNA



wild type 783-800

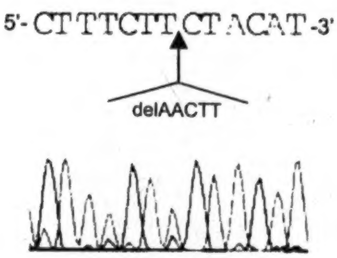


wild type 783-800

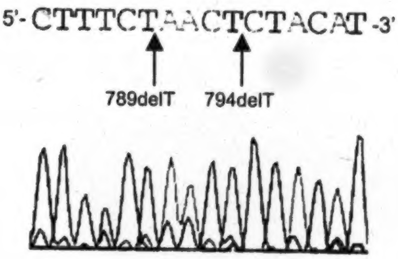


wild type 815-828

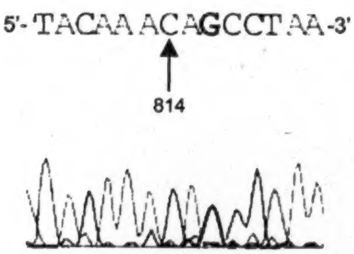
Mutant ELOVL4 cDNA



790-794delAACTT



789delT+794delT



814G→C (E272Q)

Figure 7.3 Sequence alterations affecting putative functional domains in the ELOVL4 protein introduced into wild-type ELOVL4 cDNA by site-directed mutagenesis. Electrophoretograms of portions of the wild-type cDNA sequence are given on the left and those of the engineered mutations in the corresponding locations are shown on the right. Numbers refer to nucleotide positions within the ELOVL4 cDNA sequence (amino acid position within the encoded protein is given in round parentheses). Sequences of the reverse complement of the coding strand (i.e. non-coding strand) are shown in some cases [rev].

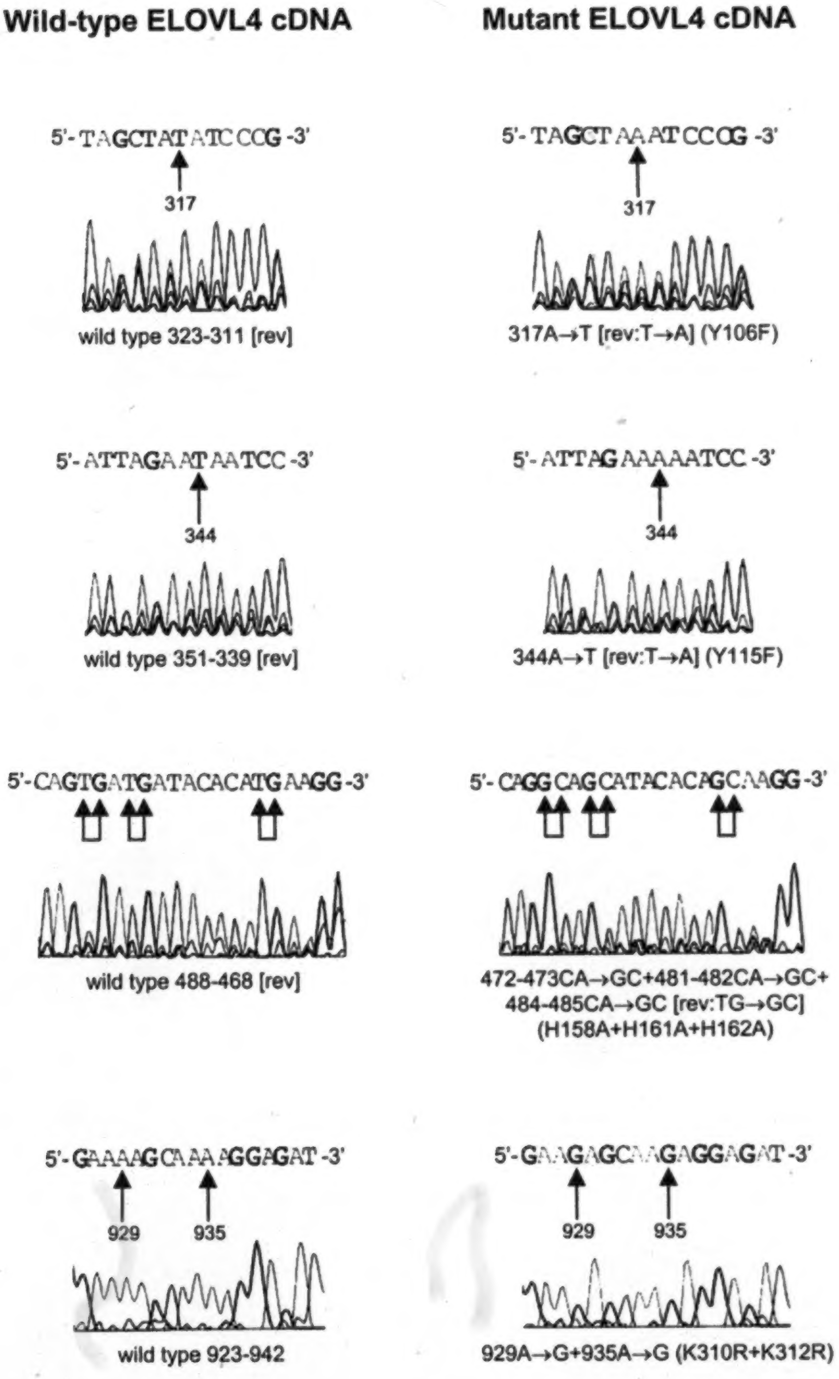

Wild-type ELOVL4 cDNA
Mutant ELOVL4 cDNA
5'- TAGCTATATCCCG -3'
317
wild type 323-311 [rev]
5'- TAGCTAAATCCCG -3'
317
317A→T [rev:T→A] (Y106F)
5'- ATTAGAATAATCC -3'
344
wild type 351-339 [rev]
5'- ATTAGAAAAATCC -3'
344
344A→T [rev:T→A] (Y115F)
5'- CAGTGATGATACACATGAAGG -3'
wild type 488-468 [rev]
5'- CAGGCAGCATACACAGCAAGG -3'
472-473CA→GC+481-482CA→GC+
484-485CA→GC [rev:TG→GC]
(H158A+H161A+H162A)
5'- GAAAAGCAAAAGGAGAT -3'
929
935
wild type 923-942
5'- GAAGAGCAAGAGGAGAT -3'
929
935
929A→G+935A→G (K310R+K312R)

Figure 7.4 ELOVL4 expression constructs for mammalian cell transfections. *A*, The pEGFP-C1 expression vector into which the wild-type and mutant ELOVL4 cDNAs were inserted for synthesis of GFP-fusion proteins in transfected cells. *B*, Schematic representation of the expression constructs consisting of the full-length open reading frame of the *ELOVL4* gene (blue) or the engineered cDNAs containing macular disease-associated mutations (blue/red) or a mutation disrupting the putative ER-retention signal (blue/white) with an in-frame amino-terminal GFP tag (green) positioned downstream of a cytomegalovirus (CMV) promoter. Sequence alterations introduced into the ELOVL4 cDNA are indicated above the corresponding mutant construct and the resulting amino acid changes in the encoded protein are given below the construct. Asterisks indicate the position of the first-encountered stop codon in the transcript. ELOVL4(del928-942/ΔKAKGD) mutant cDNA generated by Dr. Rajesh Ambasudhan (University of Michigan).

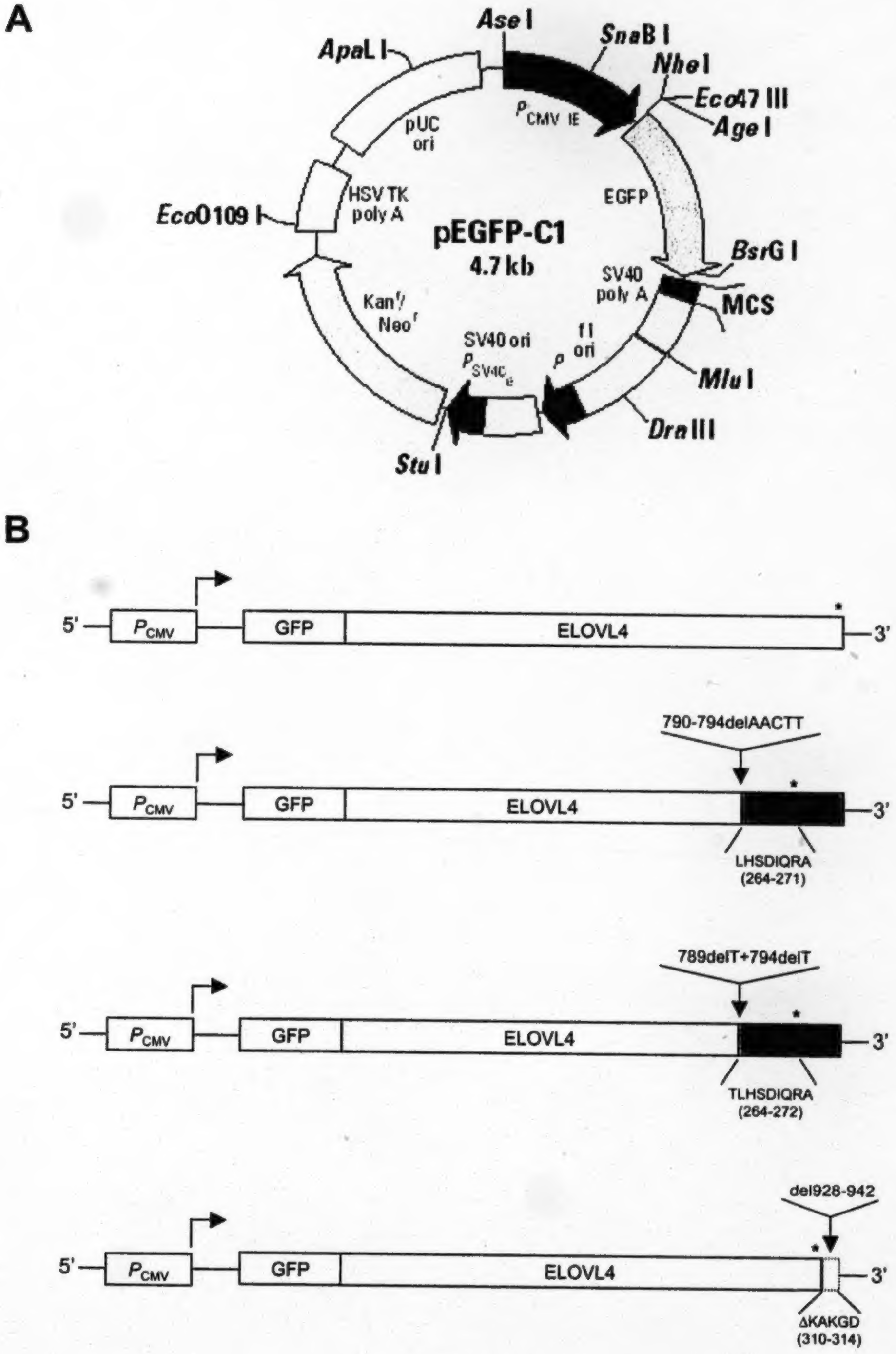
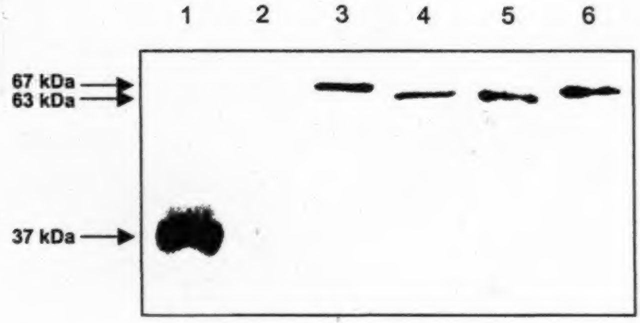

A
Ase I
SnaB I
Nhe I
Eco47 III
Age I
ApaL I
pUC ori
P CMV IE
EGFP
HSV TK poly A
Eco0109 I
pEGFP-C1
4.7 kb
BsrG I
SV40 poly A
MCS
Kan r/ Neo r
f1 ori
SV40 ori
P SV40 e
P
Mlu I
Dra III
Stu I
B
5'
P CMV
GFP
ELOVL4
*
3'
790-794delAACTT
LHSDIQRA (264-271)
789delT+794delT
TLHSDIQRA (264-272)
del928-942
ΔKAKGD (310-314)

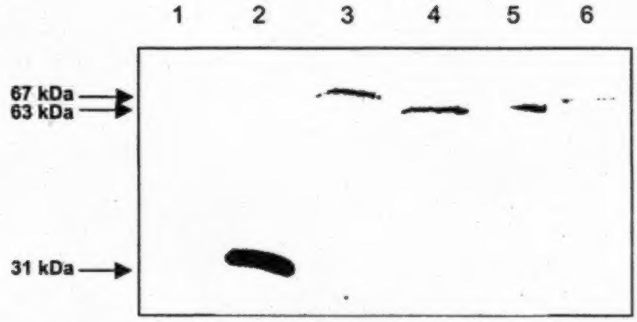
Figure 7.5 Western blot analysis of GFP-ELOVL4 fusion proteins expressed in COS-7 cells transfected with pEGFP-ELOVL4 constructs. Immobilized proteins from total cell lysates were probed with affinity-purified anti-ELOVL4 polyclonal antibody (Panel A) or anti-GFP antibody (Panel B). Control lanes in Panel A contain protein extracts from human retinal tissue (lane 1; positive control) and from COS-7 cells transfected with the empty pEGFP-C1 vector (lane 2; negative control). Control lanes in Panel B contain protein extracts from mock-transfected COS-7 cells (lane 1; negative control) and from COS-7 cells transfected with the empty pEGFP-C1 vector (lane 2; positive control). Lanes 3-6 in panels A and B contain lysates from COS-7 cells transfected with the engineered wild-type or mutant ELOVL4 cDNA constructs as follows: lane 3, pEGFP-ELOVL4 (wild-type); lane 4, pEGFP-ELOVL4 (790-794delAACTT); lane 5, pEGFP-ELOVL4 (789delT+794delT); lane 6, pEGFP-ELOVL4 (del928-942/ΔKAKGD). Positions of protein size markers are shown to the left of the blots (*kDa*, kilodaltons). Autoradiographic images provided by Dr. Rajesh Ambasudhan.

A


1
2
3
4
5
6
67 kDa
63 kDa
37 kDa

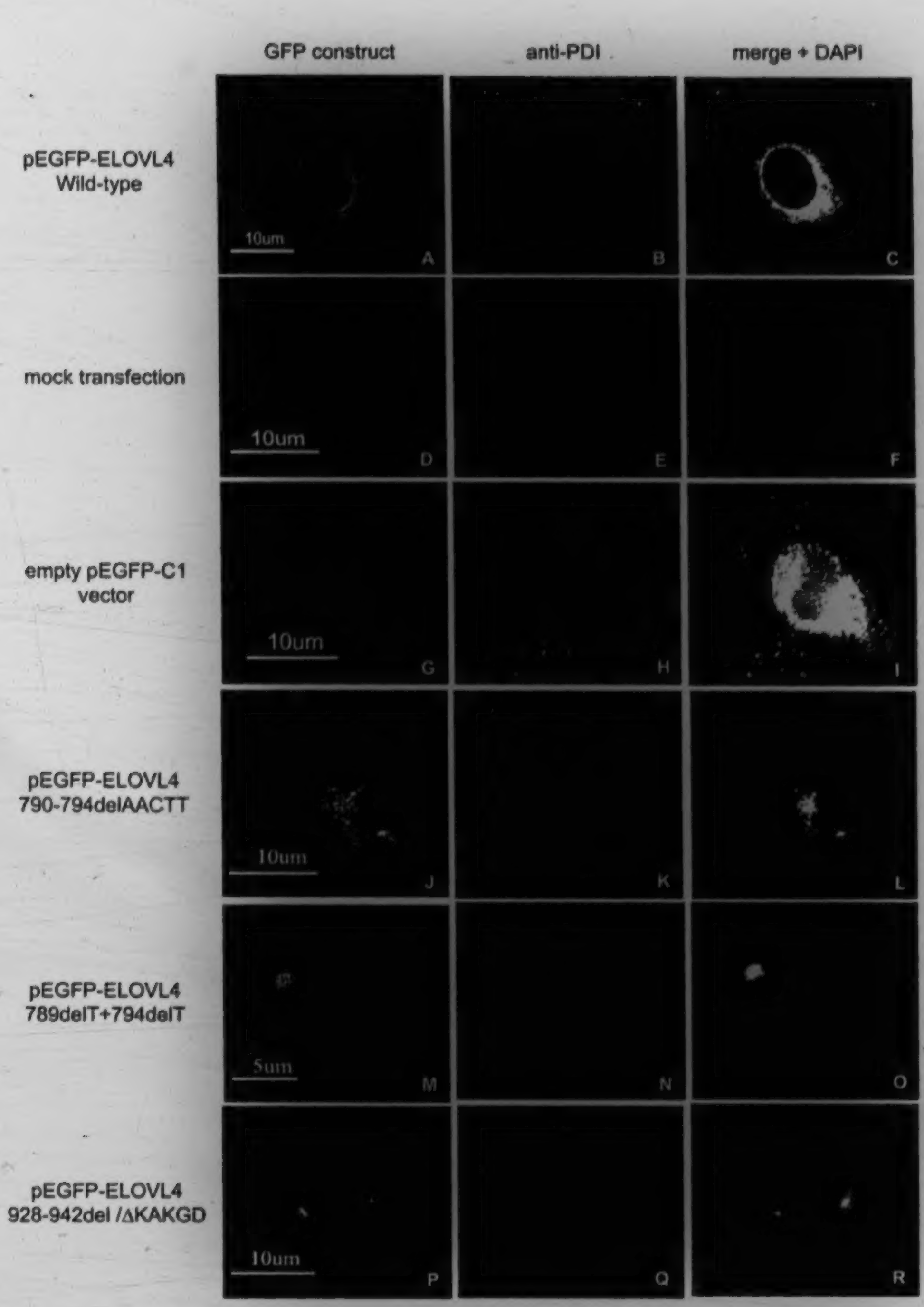
anti-ELOVL4

B


1
2
3
4
5
6
67 kDa
63 kDa
31 kDa

anti-GFP

Figure 7.6 Immunofluorescence images of COS-7 cells transfected with pEGFP-ELOVL4 cDNA constructs encoding wild-type and mutant ELOVL4 fusion proteins. Wild-type ELOVL4-overexpressing cells are shown in Panels A-C. Images from control transfections are shown in Panels D-F (mock-transfected cells) and Panels G-I (empty pEGFP-C1 vector-transfected cells). Mutant ELOVL4 cDNA constructs harbouring macular disease-associated sequence alterations are transfected into cells shown in Panels J-L (pEGFP-ELOVL4 790-794delAACTT) and Panels M-O (pEGFP-ELOVL4 789delT+794delT). Cells transfected with constructs encoding a truncated ELOVL4 protein lacking the putative carboxy-terminal ER-retention signal are shown in Panels P-R (pEGFP-ELOVL4 928-942del/ΔKAKGD). Subcellular distribution of the EGFP-ELOVL4 fusion proteins are visualized in Panels A, D, G, J, M, and P by the autofluorescence of GFP (green). Cells shown in Panels B, E, H, K, N, Q were labeled with the ER marker anti-PDI antibody and detected with a TRITC-conjugated secondary antibody (red). Images from the GFP-autofluorescent and anti-PDI/TRITC-stained panels were overlaid to generate the images in Panels C, F, I, L, O, and R (DAPI-stained cell nuclei are shown in blue). Note the near-complete co-localization of the ELOVL4 wild-type protein with the ER marker protein (Panel C, yellow merge signal), as opposed to the non-overlapping signals observed in cells expressing the mutant ELOVL4 proteins (Panels L, O, and R). Images provided by Dr. Rajesh Ambasudhan.


GFP construct
anti-PDI
merge + DAPI
pEGFP-ELOVL4 Wild-type
mock transfection
empty pEGFP-C1 vector
pEGFP-ELOVL4 790-794delAACTT
pEGFP-ELOVL4 789delT+794delT
pEGFP-ELOVL4 928-942del /ΔKAKGD
10um
A
B
C
10um
D
E
F
10um
G
H
I
10um
J
K
L
5um
M
N
O
10um
P
Q
R

Chapter 8

Summary and Conclusion

8.1 SUMMARY AND CONCLUSION

Recent years have witnessed extraordinary progress in elucidating the genetic basis of monogenic ocular disorders. Since the discovery of mutations in the human rhodopsin gene as the first determined genetic cause of inherited retinal degeneration,[1] rapid advances have been made in the mapping of numerous and varied forms of retinal disease, the identification of candidate and *bona fide* retinal disease genes, and an understanding of the biochemical mechanisms that contribute to disease pathogenesis. The work described in this thesis represents a prime example of the path of investigation followed from the initial recognition and description of a disease phenotype to determining the specific cellular events that may be responsible for the defects observed in affected individuals. In addition, the expeditious pace with which our studies have progressed demonstrates the utility and effectiveness of emphasizing phenotypic similarities versus differences, genealogical methods, recently developed human genome resources, and the collaborative process.

Cases of autosomal dominant Stargardt-like macular dystrophy (STGD3; OMIM #600110) segregating within families had been previously described,[2-6] however the combination of clinical features of the disease identified in the Canadian pedigree described in Chapter 2 appeared to represent a distinct disorder. Clinical variation among macular dystrophy patients, even within members of the same family, sometimes leads to the classification of different disorders that in fact represent simple phenotypic variability of the same disease. This has occurred for other forms of macular degeneration such as North Carolina macular dystrophy.[7,8] Thus the first enabling steps towards determining the genetic basis of retinal disease might be considered to be accurate diagnosis and the propensity to consolidate rather than isolate similar phenotypes.

We adopted this approach and began our genetic mapping studies for the Canadian STGD3 pedigree by searching for linkage to chromosome 6q, a human genomic region to which other retinal diseases had been mapped, and in particular those with similar clinical features as observed in the family under study. Through standard mapping and linkage analysis techniques, we observed positive linkage and defined a 23-cM interval on chromosome 6q to harbour the disease-causing gene and mutation for the family. By focusing our efforts on a particular region of a particular chromosome rather than initiating a genome-wide scan for linkage involving the use of widely dispersed human genetic markers, we were quickly able to establish the chromosomal location of the genetic defect underlying the Canadian STGD3 phenotype and proceed with further steps to identify the precise genetic lesion.

Continuing with our approach of uniting similar clinical entities, and limited by our access to other known family members, we sought to further reduce the disease gene-containing genetic interval by incorporating additional patients into our study. Genealogical analysis assists in elucidating whether common clinical characteristics of disease arise within individuals descended from a common founder, and who would thus be determined to share the underlying defect. The power of combining genetic data from ancestrally related families in the search for our disease gene was first demonstrated by our finding that the Canadian STGD3 pedigree was related to a family based in the United States in which autosomal dominant atrophic macular degeneration (adMD) was described. While not identical, the two disorders shared several clinical features, and our genetic mapping studies confirmed the presence of a common disease haplotype among affected individuals in both families. This established a larger pool of patient samples, and the combination of the recombinant breakpoint data from both branches of the family allowed us to reduce the STGD3/adMD interval on chromosome 6q to approximately 8 cM.

The completion of the Human Genome Project during the course of our investigation greatly influenced the methodology we adopted and the progress we subsequently made towards candidate and confirmed disease gene identification. Instead of creating a traditional physical transcript map for our genomic interval of interest, Chapter 3 describes the generation of a comprehensive *in silico* retinal and retinal pigment epithelial transcript map for the 6q12-q16 region. Development of this resource has enabled us to define several potential retinal disease genes, including *C6orf37*, as described in Chapter 4, and also the *ELOVL4* gene, as outlined in Chapter 5. Chapter 5 also details another route taken towards isolating the *ELOVL4* gene and ultimately identifying a mutation in this gene as the causative genetic defect in STGD3/adMD that was once again the result of both genealogical and bioinformatic methods. The combined genetic data derived from five independent branches of the same extended kindred further reduced the minimal genomic interval to 0.6 cM, a manageable size to begin scanning genomic sequence for potential genes. *ELOVL4* was soon discovered, and mutational analysis revealed it to be the STGD3/adMD disease gene.

As a novel gene, nothing was previously known about the function of *ELOVL4*, however bioinformatics again rapidly enabled characterization of the gene and prediction of its biological function. The results of such analysis are provided in Chapters 5, 6, and 7, and have since led to the development of assays to test the importance of the predicted functional regions of the encoded protein. The first two documented macular disease-causing mutations in *ELOVL4* result in similar protein truncation and the loss of a putative carboxy-terminal ER-targeting signal. We have tested the effects of these mutations on the subcellular distribution of the ELOVL4 protein and the suspected defect in localization was verified in a mammalian cell culture model system, as presented in Chapter 7. Protein mislocalization has also very recently been reported for a third *ELOVL4* mutation encoding a Y270stop allele in Dutch macular dystrophy patients.[9] These studies provide experimental evidence for the possible cellular role of this protein as an enzyme involved in fatty acid elongation that takes place in the ER. Additional mutations introduced into the *ELOVL4* coding region will enable us to test the significance of other predicted protein domains. This can be accomplished using the described cell culture system for either additional localization studies or for *in vitro* fatty acid elongation assays performed using whole cells or microsomes prepared from the transfected cells, as have been used to test the function of other components of the fatty acid elongation pathway.[10-12] We have also begun to develop a functional assay in the yeast *Saccharomyces cerevisiae* for the evaluation of both wild-type and mutant ELOVL4 protein activities on the elongation of fatty acid substrates,[13] as has been described for a number of ELOVL4 homologues identified from a wide range of species.[14-17] Such studies will be able to definitively establish whether the *ELOVL4* gene product is in fact a fatty acid elongase as is presumed based on its amino acid sequence characteristics. This will be critical to understanding how mutations in the gene may lead to retinal cell dysfunction that causes degeneration of the macula in affected patients. A transgenic mouse model has also recently been documented in which a mutant version of the human *ELOVL4* gene is expressed.[18] This provides another tool for analysis of *ELOVL4* gene and mutation function *in vivo*.

If *ELOVL4* does prove to encode a fatty acid biosynthetic enzyme, this would classify STGD3/adMD as an inborn error of lipid metabolism and the first retinal disorder to be defined as such. The pathogenesis of such diseases is typically attributed to either the deficiency of a metabolite normally synthesized by the wild-type enzyme and/or the accumulation and possible aberrant metabolism of substrate molecules. This in turn would suggest possible therapeutic strategies for macular dystrophy patients arising from *ELOVL4* gene mutation. It has been speculated that the product of ELOVL4 enzyme action might be the abundant retinal fatty acid, DHA. Given the importance of DHA to the structure and function of the retina, the implications of DHA deficiency on visual performance, and the recognition that the process of fatty acid elongation of precursor lipids is required for its synthesis,[19] an obvious course of therapy for STGD3/adMD patients might be DHA supplementation. Moreover, a recent report describes the potential influence of dietary fat intake on disease manifestation in macular dystrophy patients harbouring an *ELOVL4* mutation, with an inverse correlation observed between DHA levels in the blood and phenotypic severity.[20] Studies of the effects of DHA supplementation on retinal function have been initiated for the Canadian STGD3 family,[21,22] and while modest effects have been observed in some individuals, additional testing through controlled clinical trials is required to accurately assess the efficacy of this potential treatment. Confirmation of ELOVL4 activity as a fatty acid elongase and characterization of its physiological substrates and products will further assist in the development of applicable treatment modalities for visual impairment caused by ELOVL4 dysfunction.

The results of our work have implications for each step of the process we have followed. The diagnosis of individuals in the Canadian STGD3 family, related family members, and other macular degeneration patients with a similar clinical presentation has been facilitated by the ability to test for specific mutations in a particular gene. Continued characterization of retinal genes on chromosome 6q will contribute to an expanded knowledge of retinal biology and may uncover additional disease genes for disorders of the retina that map to the characterized genomic interval. Additional studies of *ELOVL4* gene function will provide insight into biochemical processes occurring in the retina and molecular mechanisms of pathogenesis leading to photoreceptor cell dysfunction and death. Knowing the precise biological role of the *ELOVL4* gene will further recommend possible modes of treatment in macular dystrophy patients, potentially validating those that have already been initiated, and perhaps suggesting other courses of therapy. We have helped to pave the way towards further understanding and hopefully one day overcoming the devastation of retinal degeneration and consequent vision loss.

8.2 REFERENCES

1. Dryja TP, McGee TL, Reichel E, Hahn LB, Cowley GS, Yandell DW, Sandberg MA, Berson EL. (1990) A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. Nature 343:364-366.

2. Cibis GW, Morey M, Harris DJ. (1980) Dominantly inherited macular dystrophy with flecks (Stargardt). *Arch. Ophthalmol.* 98:1785-1789.

3. Lopez PF, Maumenee IH, de la Cruz Z, Green WR. (1990) Autosomal-dominant fundus flavimaculatus: clinicopathologic correlation. *Ophthalmology* 97:798-809.

4. Mansour AM. (1992) Long-term follow-up of dominant macular dystrophy with flecks (Stargardt). *Ophthalmologica* 205:138-143.

5. Stone EM, Nichols BE, Kimura AE, Weingeist TA, Drack A, Sheffield VC. (1994) Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. *Arch. Ophthalmol.* 112:765-772.

6. Zhang K, Bither PP, Park R, Donoso LA, Seidman JG, Seidman CE. (1994) A dominant Stargardt's macular dystrophy locus maps to chromosome 13q34. *Arch. Ophthalmol.* 112:759-764.

7. Small KW, Hermsen V, Gurney N, Fetkenhour CL, Folk JC. (1992) North Carolina macular dystrophy and central areolar pigment epithelial dystrophy: one family, one disease. *Arch. Ophthalmol.* 110:515-518.

8. Small KW. (1998) North Carolina macular dystrophy: clinical features, genealogy, and genetic linkage analysis. *Trans. Am. Ophthalmol. Soc.* 96:925-961.

9. Karan G, Maugeri A, Yang Z, Jiang L, Li X, Cremers F, Zhang K. (2004) Functional consequences of a novel mutation in *ELOVL4* causing macular dystrophy. *Invest. Ophthalmol. Vis. Sci.* 45:E-Abstract 1766.

10. Leonard AE, Kelder B, Bobik EG, Chuang LT, Lewis CJ, Kopchick JJ, Mukerji P, Huang YS. (2002) Identification and expression of mammalian long-chain PUFA elongation enzymes. *Lipids* 37:733-740.

11. Moon YA, Shah NA, Mohapatra S, Warrington JA, Horton JD. (2001) Identification of a mammalian long chain fatty acyl elongase regulated by sterol regulatory element-binding proteins. *J. Biol. Chem.* 276:45358-45366.

12. Moon YA, Horton JD. (2003) Identification of two mammalian reductases involved in the two-carbon fatty acyl elongation cascade. *J. Biol. Chem.* 278:7335-7343.

13. Lagali PS, Ambasudhan R, Liu J, Ayyagari R, Wong, PW. (2003) Use of the yeast *Saccharomyces cerevisiae* as a model organism to study human retinal disease gene mutations. *Invest. Ophthalmol. Vis. Sci.* 44:E-Abstract 5097.

14. Parker-Barnes JM, Das T, Bobik E, Leonard AE, Thurmond JM, Chaung LT, Huang YS, Mukerji P. (2000) Identification and characterization of an enzyme involved in the elongation of n-6 and n-3 polyunsaturated fatty acids. *Proc. Natl. Acad. Sci. U. S. A.* 97:8284-8289.

15. Leonard AE, Bobik EG, Dorado J, Kroeger PE, Chuang LT, Thurmond JM, Parker-Barnes JM, Das T, Huang YS, Mukerji P. (2000) Cloning of a human cDNA encoding a novel enzyme involved in the elongation of long-chain polyunsaturated fatty acids. *Biochem. J.* 350:765-770.

16. Das T, Thurmond JM, Bobik E, Leonard AE, Parker-Barnes JM, Huang YS, Mukerji P. (2000) Polyunsaturated fatty acid-specific elongation enzymes. *Biochem. Soc. Trans.* 28:658-660.

17. Zank TK, Zahringer U, Lerchl J, Heinz E. (2000) Cloning and functional expression of the first plant fatty acid elongase specific for Δ^6-polyunsaturated fatty acids. *Biochem. Soc. Trans.* 28:654-658.

18. Locke KG, Karan G, Yang Z, Zhang K, Birch DG. (2004) Human mutant ELOVL4 causes retinal dystrophy in transgenic mice. *Invest. Ophthalmol. Vis. Sci.* 45:E-Abstract 1769.

19. Leonard AE, Pereira SL, Sprecher H, Huang YS. (2004) Elongation of long-chain fatty acids. *Prog. Lipid Res.* 43:36-54.

20. Hubbard AF, Askew EW, Thompson CJ, Singh N, Leppert M, Bernstein PS. (2004) Association of adipose and red blood cell lipids with severity of dominant Stargardt macular dystrophy. *Invest. Ophthalmol. Vis. Sci.* 45:E-Abstract 1770.

21. MacDonald IM, Hebert M, Yau RJ, Flynn S, Jumpsen J, Suh M, Clandinin MT. (2004) Effect of docosahexaenoic acid supplementation on retinal function in a patient with autosomal dominant Stargardt-like retinal dystrophy. *Br. J. Ophthalmol.* 88:305-306.

22. MacDonald IM, Flynn SJ, Lee TKM, Yau RB, Jumpsen J, Suh M, Clandinin MT, Hebert M. (2004) Docosahexaenoic acid supplementation in a family with a mutation in *ELOVL4. Invest. Ophthalmol. Vis. Sci.* 45:E-Abstract 1768.